The Dissertation Committee for Lin Hong
Certifies that this is the approved version of the following dissertation:

BIOCHEMICAL STUDIES OF THE ENZYMES INVOLVED IN DEOXYSUGAR D-FOROSAMINE BIOSYNTHESIS

Committee:

_____

Hung-wen Liu, Supervisor

_____

Walter Fast

_____

David W. Hoffman

_____

Sean M. Kerwin

_____

Jon D. Robertus

BIOCHEMICAL STUDIES OF THE ENZYMES INVOLVED IN DEOXYSUGAR D-FOROSAMINE BIOSYNTHESIS

by

LIN HONG, B. S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2004

..

BIOCHEMICAL STUDIES OF THE ENZYMES INVOLVED IN DEOXYSUGAR D-FOROSAMINE BIOSYNTHESIS

Publication No. _____

Lin Hong, Ph.D.
The University of Texas at Austin, 2004

Supervisor: Hung-wen Liu

Deoxysugars are indispensable structural components of many biologically active natural products and are essential for many significant cellular processes. In this work, the biosynthetic pathway of TDP-forosamine, which is a 2,3,4,6-tetradeoxy sugar component of the potent, and environmentally benign insecticide spinosyn, has been established without ambiguity. Five genes, *spnO*, *spnN*, *spnQ*, *spnR*, and *spnS* have been cloned from *Saccharopolyspora spinosa* chromosomal DNA and expressed in *E. coli*. The encoded proteins have been purified and characterized by *in vitro* assays. The products of four reactions have been isolated and characterized with MS and/or NMR. SpnO was demonstrated to be a 2,3-dehydratase, while SpnN was shown to be a 3-

ketoreductase, catalyzing equatorial hydroxyl formation at C-3. Studies of the combined action of SpnO and SpnN reinforce the previous finding that C-2 deoxygenation is accomplished through a β-elimination/reduction type mechanism. SpnQ is a homologue of the well-characterized CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase (E_1) which in combination with CDP-6-deoxy-4-keto-$\Delta^{3,4}$-glucoseen reductase (E_3) catalyzes the C-O bond cleavage at C-3 in the biosynthesis of CDP-ascarylose in *Yersinia pseudotuberculosis*. In the biosynthesis of TDP-forosamine, SpnQ was shown to catalyze an analogous C-3 deoxygenation but instead using ferredoxin/ferredoxin reductase pair as the electron transfer media. Additionally, in the absence of ferredoxin/ferredoxin reductase to allow C-3 deoxygenation, and in the presence of L-glutamate as an amino donor, SpnQ was shown to act as a 4-aminotransferase, converting TDP-2,6-dideoxy-4-keto-D-glucose to TDP-4-amino-2,4,6-trideoxy-D-glucose. Due to the instability of 2-deoxysugars, an analog of TDP-4-amino-2,3,4,6-tetradeoxy-glucose was used to characterize the late steps in the pathway. SpnR was shown to be a 4-aminotransferase which, interestingly, can recognize both the SpnQ product and the SpnN product as substrates. SpnS was shown to be a 4-amino-*N*,*N*-dimethyltransferase. Mechanistic investigation revealed that both *N*-methyl groups of TDP-forosamine are installed by SpnS in a step-wise manner via a monomethylated intermediate. Together, studies described in this thesis have extensively enriched our knowledge of the enzymes involved in deoxysugar biosynthesis. The newly discovered substrate flexibility of TDP-forosamine pathway enzymes will also be very useful in generating engineered sugar biosynthetic pathway for use in generating novel natural products.

TABLE OF CONTENTS

Chapter 2. Characterization of SpnO, a TDP-6-deoxy-4-keto-D-glucose 2,3-dehydratase and SpnN, a TDP-2,6-dideoxy-3,4-diketo-D-glucose 3-ketoreductase

LIST OF FIGURES

LIST OF SCHEMES

LIST OF EQUATIONS

LIST OF TABLES

LIST OF ABBREVIATIONS

$A_{280}/A_{340}/A_{562}/A_{410}$	absorbance at 280nm/340nm/562nm/410nm
ADP	adenosine diphosphate
CDP	cytidine diphosphate
DEAE	diethylaminoethyl
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
TDP	thymidine diphosphate
e	electron(s)
E_{od}	NDP-D-glucose 4,6-dehydratase
E_1	CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase
E_3	CDP-6-deoxy-4-keto-$\Delta^{3,4}$-glucoseen reductase
EPR	electron paramagnetic resonance
$FAD/FADH_2$	flavin adenine dinucleotide, oxidized/reduced form
[2Fe-2S]	two iron, two sulfur redox center
[3Fe-4S]	three iron, four sulfur redox center
[4Fe-4S]	four iron, four sulfur redox center
FLD	flavodoxin
FNR	flavodoxin reductase
FPLC	fast protein liquid chromatography
Fru-6-P	fructose 6-phosphate
Gal	D-galactose
GalNAc	*N*-acetyl-D-galactosamine
Glc	D-glucose
GlcNAc	*N*-acetyl-D-glucosamine
GlmS	glucosamine-6-phosphate synthase

GDP	guanine diphosphate
h	hour(s)
HPLC	high performance liquid chromatography
IPTG	isopropylthio-β-galactoside
kb	kilo base pairs
kDa	kilo Dalton
α-KG	α-ketoglutarate
LB	Luria-Bertani medium
LPS	lipopolysaccharide
μ	micro
m	milli
M	molar
NDP	nucleoside diphosphate
NAD^+/NADH	nicotinamide adenine dinucleotide, oxidized/ reduced forms
$NADP^+$/NADPH	nicotinamide adenine dinucleotide phosphate oxidized/reduced forms
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
orf/ORF	open reading frame
PCR	polymerase chain reaction
PKS	polyketide synthase
PMP	pyridoxamine-5'-phosphate
PLP	pyridoxal-5'-phosphate
RBS	ribosome binding site
RNA	ribonucleic acid
RNR	ribonucleotide reductase
s	second(s)
SAM	*S*-adenosyl methionine

SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel Electrophoresis
S_N2	bimolecular nucleophilic displacement reaction
TB	terrific broth
TE	thioesterase
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
UDP	uridine diphosphate

Chapter 1. Background and Significance

1. INTRODUCTION

Carbohydrates, proteins and nucleic acids are three important constituents of biological systems. Although each has its respective role in providing energy for the living organism, catalyzing various types of biological reactions, and serving as the template for the reproduction and evolution of organisms, they also serve specific functions by acting in combination.[1-7] For example, glycopeptides are indispensable components of cell wall structures. It is also well known that the antigenic determinants of the ABO(H) blood group and the related lewis blood group are carbohydrates attached to a peptide.[8,9] The ribosome that is essential in protein production by translating mRNA is an assortment of nucleic acids and proteins. The universal presence of carbohydrates in living systems has inspired extensive research to explore their diverse functions in microorganisms, plants, and animals.[10-16]

Deoxysugars are an important class of carbohydrates that occur widely in nature. They are commonly defined as monosaccharides with one or more hydroxyl group(s) replaced by hydrogen(s). However, since deoxysugars often contain functional groups other than hydrogens in place of the hydroxyls, they are more often defined as sugars in which one or more of the hydroxyl groups is replaced by a hydrogen (deoxysugars) or other heteroatom or heteroatomic group, such as sulfur (thiosugars), amino group (aminosugars) or nitro group (nitrosugar). There is also a group of deoxysugars where an additional alkyl-branched chain is attached to a backbone carbon on the sugar ring.[17-19] These are referred to as alkyl-branched chain deoxysugars. Deoxysugars can be found as

elements of lipopolysaccharides (LPS), glycoproteins, and a great variety of secondary metabolites with biological activities.

LPS is a constituent of the gram-negative bacterial cell wall and is known to be responsible for the endotoxic effect of these bacteria on the host during infection.[20-23] LPS has a common architecture consisting of a lipid to which a carbohydrate portion is covalently linked (Figure 1-1). This carbohydrate portion contains an oligosaccharide

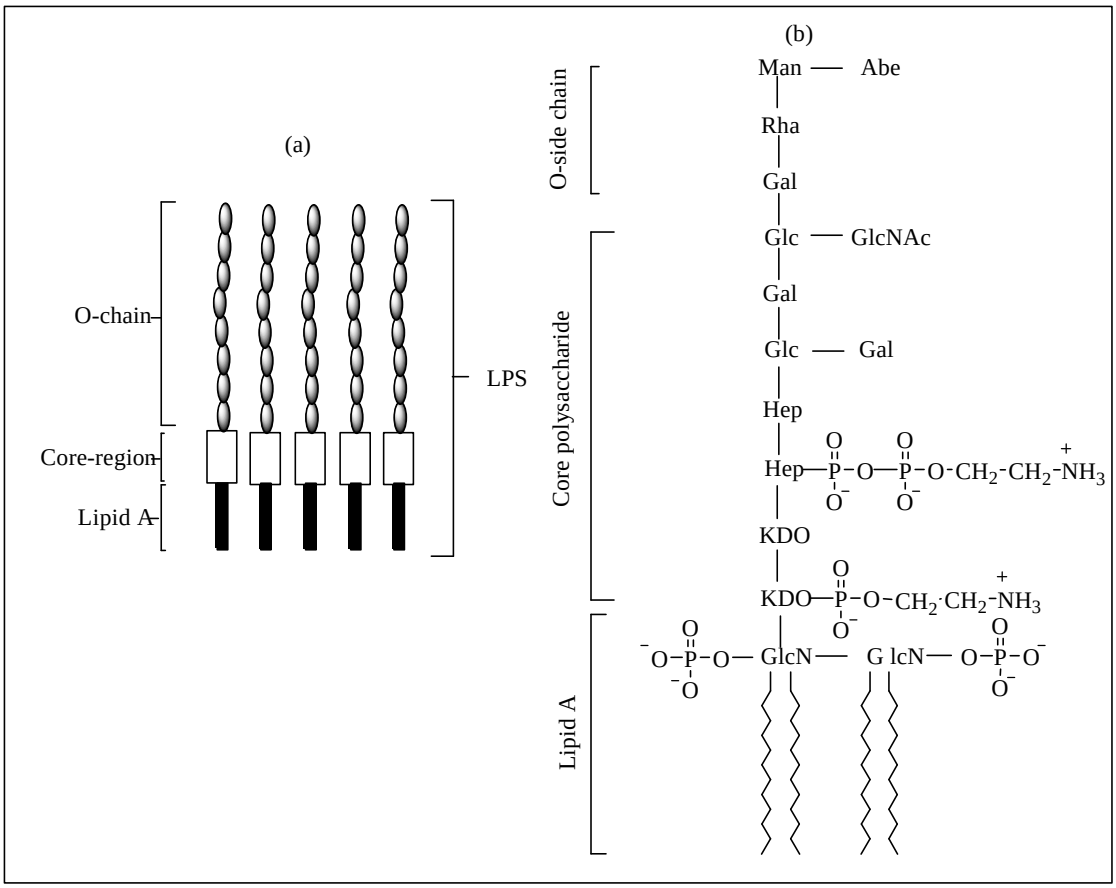


Figure 1-1. (a) General tripartite structure of LPS. (b) LPS of *Salmonella typhimurium*. Abbreviations: Abeq, abequose; Gal, galactose; GlcN, *N*-acetylglucosamine; Glu, gluocse; Hep, L-glycero-D-mannoheptose; KDO, keto-deoxy-octanoate; Man, mannose. Adapted from Nat. Prod. Rep. 1999, 16, 283-299. Scrimgeour, M. Biochemistry; Prentice Hall:New Jersey, 1994.

unit with up to 15 monosaccharides called the core region, and a polysaccharide extension made of repeating units referred to as the *O*-chain.[24,25,31] The structures of LPS are similar to some glycoproteins that also have a tripartite structural organization, where the usually conserved core regions serves as a linkage connecting a glycan chain to a polypeptide.[26,27] Shown in Figure 1-2 are two structures of *N*-linked glycopeptides having the tripartite organization. While the

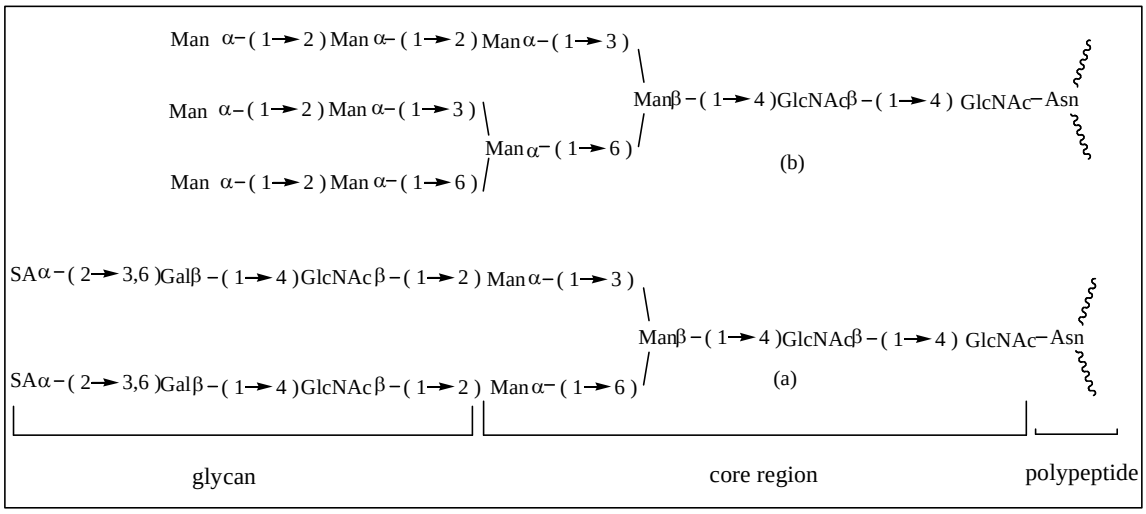


Figure 1-2. Structures of N-linked oligosaccharides. (a) High-mannose chain. (b) Complex chain. Scrimgeour, M. Biochemistry; Prentice Hall:New Jersey, 1994.

core region of the LPS in most bacteria is very similar, the *O*-chains display a broad structural diversity among different organisms.[24,34] It has been well documented that the *O*-chains of the LPS are rich in deoxysugars which are directly involved in host recognition, modulation of bacterial virulence, and cell-cell adhesion.[21,28,29,32,35,36,37] In

addition, studies of the structure and function of LPS have shown that the immunospecificity of the pathogenic bacteria is in large part determined by the deoxysugar unit at the non-reducing end of the *O*-chain.[2,30,38,39]

Deoxysugars are also found as structural components in naturally occurring antimicrobial agents and anticancer drugs. For example, macrolide antibiotics are made up of a cyclic polyketide aglycone with one or more sugars that are linked to the aglycone. Various deoxysugars have been found in macrolide antibiotic structures. Some of these sugars have further modifications, such as *O*-, *N*-, and *C*-methylations and acetylations. Accumulated evidence showed that deoxysugars are essential for the biological activities and the specificities of antibiotics of which they are a part.[12,39] High-resolution X-ray crystal structures of the 50S ribosomal subunit of the eubacterium *Deinococcus radiodurans* complexed with erythromycin, clarithromycin and roxithromycin have unambiguously established the interactions between the deoxysugar portions of these compounds and the nucleotides of 23S ribosomal RNA. These interactions inhibit the peptidyl transferase activity of the ribosome by interfering with the proper positioning and movement of tRNAs in the peptidyl transferase cavity.[40,41]

In light of the above structural as well as biochemical evidence, it is expected that the activity and specificity of these antibiotics could be fine-tuned by making changes to deoxysugar structures. Much efforts have been devoted to prepare derivatives of known antibiotics carrying varied deoxysugar components in order to overcome the increasingly serious drug resistance problem.[42,43,44] Although chemical synthesis has traditionally been the approach of choice, biosynthetic engineering has emerged as an alternative

means for generating structural variations in recent years. This approach is made possible by the availability of the sequencing information of a large number of deoxysugar-containing secondary metabolite biosynthetic clusters. *In vivo* and *in vitro* studies have shown that many glycosyltransferases catalyzing the attachment of deoxysugars to acceptor substrates have some degree of substrate flexibility, and thus are capable of accepting some unnatural deoxysugars generated by reengineering the deoxysugar biosynthetic machineries. It is encouraging that some of the resulting antibiotic analogs exhibit reasonable antibiotic activities.[45-51] The success of such pathway reengineering strategy depends on a thorough understanding of the functions of the enzymes encoded by each individual genes in the deoxysugar biosynthetic clusters. The studies of our laboratory have provided much insight into the enzymatic logic behind TDP-deoxysugar formation.[52,53,54] Detailed investigation of the enzymatic mechanism of the C-O bond cleavage in the biosynthesis of 3,6-dideoxyhexoses represents one of the first mechanistic studies of deoxysugar biosynthesis.[55-68] Over the years, the enzymatic mechanisms of C-O bond cleavage at other positions of the hexose ring, and the mechanisms of the C-N as well as C-C bond formation involved in amino and branched-chain sugar biosynthesis, have also been explored.[71,72,73,80,81] These endeavors not only provide valuable information on the characteristics of the enzymes involved, but also set the stage for further development of creating deoxysugar-containing bio-active compounds either through *in vivo* engineering or *in vitro* enzymatic synthesis.

2. BIOSYNTHETIC PATHWAYS OF DEOXYSUGARS

Polyketide Synthase (PKS)

Many natural products are polyketide derivatives, having a highly functionalized aglycone which is further modified at various positions by deoxysugars. According to the structures of the polyketide aglycones, these natural products can be roughly divided into three families:[82,83] the aromatic tetracyclic family,[84,85] the macrocyclic lactone family,[86] and the aminocoumarin family.[89,90] The polyketide aglycones are synthesized by polyketide synthases, which are large multifunctional enzyme complexes. There are two main types of polyketide synthases: modular polyketide synthase and iterative polyketide synthase.[83,91] Modular polyketide synthases, also known as type-I PKS, are large multienzyme complexes that synthesize macrocyclic polyketides by successive

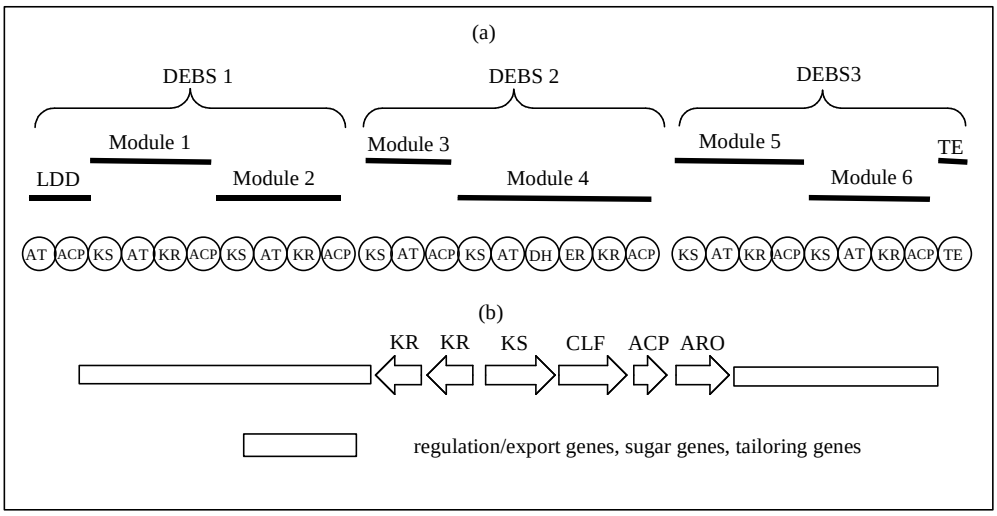


Figure 1-3 (a) Biosynthesis of 6-deoxyerythronolide B (6dEB) by deoxyerythronolide B synthase , a type I PKS. Abbreviations: KS, β -ketosynthase; AT, acyl transferase; ACP, acyl carrier protein; KR, keto reductase; DH, dehydrase; ER, enoy reductase. (b) Organisation of the gra polyketide synthase gene cluster, a type II PKS. Abbreviations: KS, keto-acyl synthase; CLF, chain length factor; ACP, acyl carrier protein; ARO, aromatase; KR, ketoreductase. Adapted from Hans, M. et al. J. Am. Chem. Soc. 2003, 125, 5366.

condensation of acetate, propionate or butyrate units. The extent of β-carbonyl-group reduction in each condensation cycle varies. Iterative PKS, also known as type-II PKS, construct polyketides mainly from acetate units, and the carbonyl group formed after each condensation cycle are usually left unreduced. The erythromycin PKS is an example of a type-I PKS, while the granaticin PKS is an example of a type-II PKS (Figure 1-3). For type-I PKS (Figure 1-4), the synthesis of a polyketide is completed by

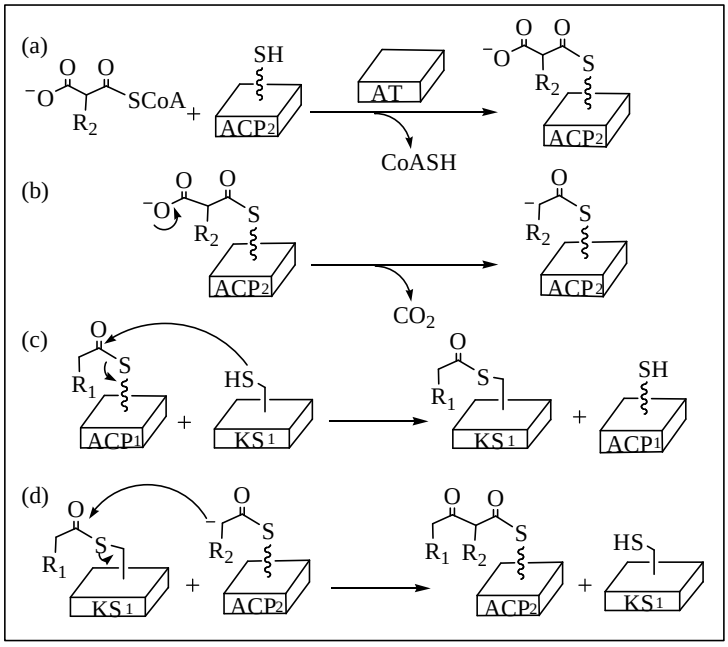


Figure 1-4. Four stages to the C-C bond-forming chain elongation steps of type I PKS assemblies: (a) malonyl- or methylmalonyl-S-ACP formation; (b) decarboxylation to generate the carbon nucleophile; (c) the acyl-S-Cys-KS donor; (d) the product, β -keto-acyl-S-ACP. Adapted from Walsh, C. T. Antibiotics: Action, Origins, Resistance. ASM Press: Washington, DC. 2003

the sequential catalysis of several modules, each of which is composed of several individual domains. The start unit is first linked to the ACP of the loading domain (module). By the catalysis of ketosynthase domain and acyl-transferase domain on the

following module, the unit is transferred to the ACP domain of the second module and the size of the chain increases by either a two or three-carbon unit, depending on the condensation unit used. The polyketide chain can be further modified by the domain in the same module, such as ketoreductase domain, dehydratase domain and enoyl reductase domain before being transferred to the next module. Each module in the type-I PKS has the basic duty of increasing the length of the polyketide chain, possibly reducing the most recently added unit, and finally passing it to the next module. When the polyketide chain is loaded onto the last module, the linear chain is cyclized and released from the PKS. As for type-II PKS (Figure 1-5), the domains are dispersed among the metabolite biosynthesis cluster. The enzymatic function of each domain is used repeatedly until the final product is formed. In each chain elongation cycle, malonyl or methylmalonyl unit is loaded onto the ACP domain, while the chain with increasing size is retained on the KS_α

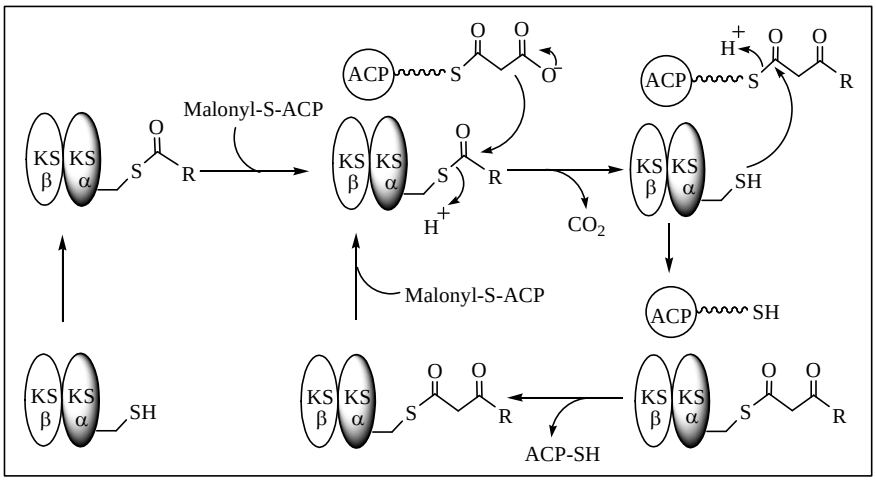


Figure 1-5. The iterative chain elongation cycles of tyep II PKS. Adapted from Walsh, C. T. Antibiotics: Actions, Origins, Resistance; ASM press: Washington, D.C. 2003

domain. The adjacently localized KS_β domain is supposed to control the length of the polyketide chain and the number of the elongation cycle. Since the work in this dissertation is not related to polyketide synthesis, the PKS-catalyzed reactions are only briefly described here. More detailed account on catalysis by each type of PKS can be found in many reviews.[83,87,88,92,93]

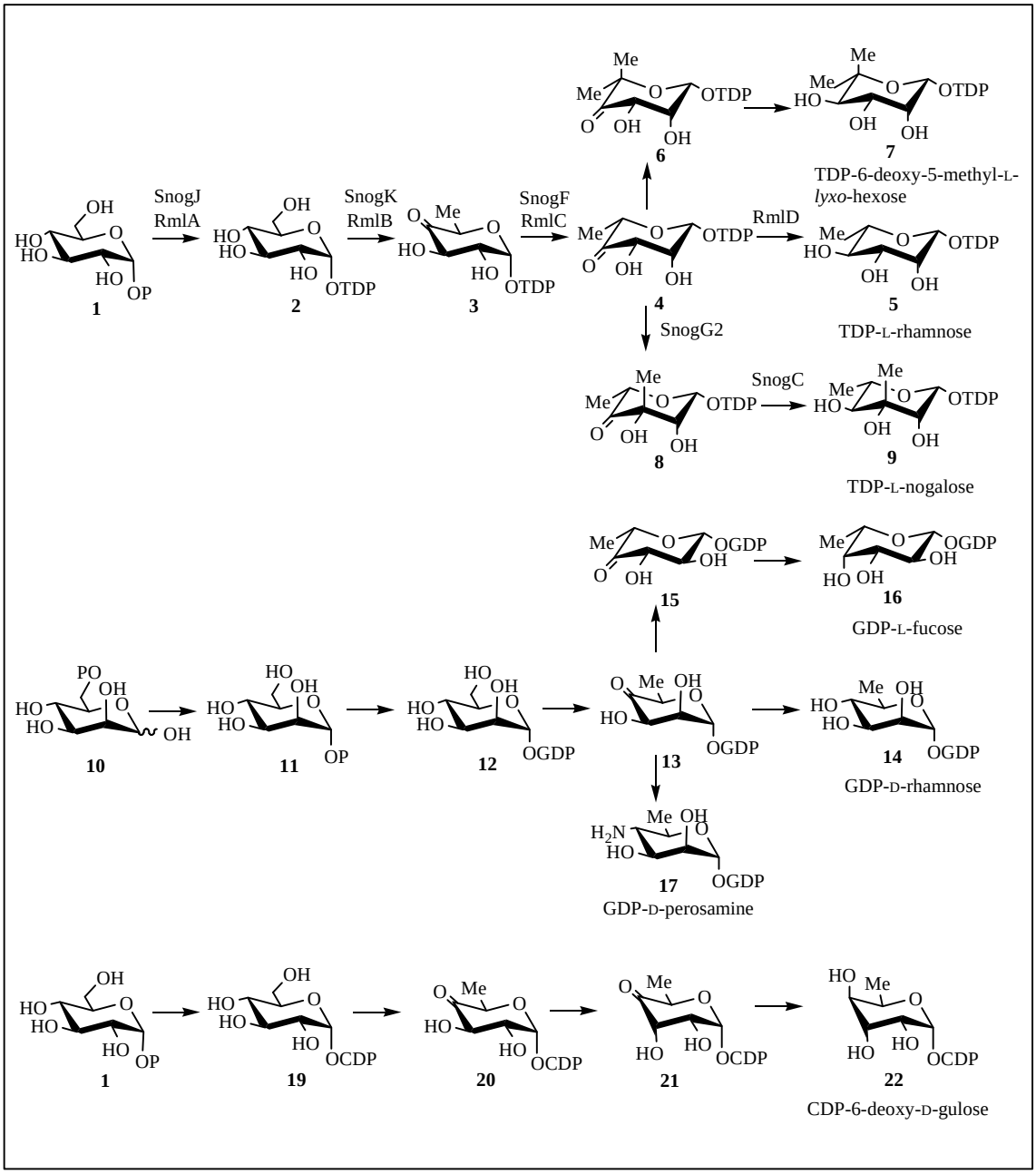
Genes in Deoxysugar Biosynthesis

The genes for the biosynthesis of deoxysugars that are attached to polyketide-derived aglycone are usually located adjacent to the PKS genes responsible for the formation of the alycone to which the sugar(s) is attached. When a sugar unit is not only a component of a secondary metabolite, but also of LPS of the producing organism, the sugar biosynthesis genes may be found outside the secondary metabolite gene cluster.[94] To synthesize such secondary metabolites, the sugar unit probably is obtained from the cellular pool. There are also cases where a sugar biosynthetic gene, whose encoded protein catalyzes a reaction common to several deoxysugar pathways that occur in the same organism, is present in only one copy. The encoded enzyme is presumably used by all pathways for which it is needed. For example, TylX3 is the dehydratase in the biosynthesis of three deoxysugars: mycaminose, mycarose, and deoxyallose, all of which are the sugar components of antibiotic tylosin. In tylosin biosynthesis gene cluster, only one copy of *tylX3* can be identified. Although the structures of the polyketides to which deoxysugars are attached and those of the O-chain of LPS which have deoxysugars as components are very different, so as the organization of the biosynthetic genes in each

pathway, the same sugar component of different natural products generally have the same biosynthetic pathway. In the following section, the biosynthetic pathways of deoxyhexoses will be illustrated. The deoxyhexoses discussed are subdivided into following groups: 6-deoxysugars, 2,6-dideoxysugars, 3,6-dideoxysugars, 2,3,6-trideoxysugars and 3,4,6-trideoxysugars. The discussion is not intended to be a complete description of deoxysugar biosynthesis, but is instead an illustration of the diversity of deoxysugar structures found in nature and the presentation of several better-studied biosynthetic pathways. The names of some enzymes are not given. This is either because the enzyme has not been assigned a name, or because the deoxysugar is made by a few different organisms, each containing a biosynthetic machinery for that dexoysugar. In most cases, the names used by different research groups have not been standardized.

Biosynthetic Pathways of 6-Deoxysugars (Scheme 1-1)

One of the best-known 6-deoxysugars, L-rhamnose, is found widely in bacteria and plants and is a common component of the cell wall and capsule of many pathogenic bacteria. The cell wall and capsule of the pathogen interact with the host during infection and help the bacteria to survive. L-Rhamnose is also a component of several secondary metabolites produced by microorganisms.[94] The precursor of rhamnose is glucose with a nucleotide diphosphate group attached at C-1. Having such a diphospho- linkage facilitates the bond formation between a nucleophile and the C-1 of the sugar since nucleotide diphosphate is a better leaving group than a hydroxyl group. In fact, in all deoxysugar biosynthesis studied so far, the sugar unit is always activated to NDP-sugar. It has been shown that the first enzyme in the TDP-L-rhamnose (**5**) biosynthetic pathway



Scheme 1- 1. 6-deoxy hexoses biosynthetic pathways.

in *Salmonella enterica* is RmlA, which is a α-D-glucose-1-phosphate thymidylyltransferase catalyzing the transfer of thymidylmonophosphate (TMP) from thymidyltriphosphate (TTP) to glucose-1-phosphate (**1**). The second enzyme, RmlB, is a

TDP-glucose-4,6-dehydratase. The product formed, TDP-6-deoxy-4-keto-glucose (**3**) is an intermediate common to many deoxysugar biosynthetic pathways and is the branch point at which various pathway diverge. The third enzyme, RmlC, catalyzes both the C-3 and C-5 epimerization reactions. Finally, RmlD reduces the 4-keto functional group to generate the final product, TDP-L-rhamnose (**5**).[24,95]

Another 6-deoxysugar, D-rhamnose, is primarily utilized in the cell wall synthesis by *Pseudomonas aeruginosa*.[24,96] Unlike L-rhamnose, D-rhamnose is made from mannose-1-phosphate (**11**). After guanidylylmonophosphate nucleotide transfer, GDP-mannose (**12**) is formed. A TDP-glucose-4,6-dehydratase homologue then catalyzes the formation of GDP-6-deoxy-4-keto-mannose (**13**). Finally, a ketoreductase reduces the C-4 position to complete the synthesis of GDP-D-rhamnose (**14**).

L-Fucose is a sugar commonly found as component of LPS and surface antigen of *E. coli*. The biosynthetic pathway of GDP-L-fucose (**16**) begins with mannose-1-phosphate (**11**), which is converted to GDP-mannose (**12**), then GDP-6-deoxy-4-keto-mannose (**13**) as in GDP-D-rhamnose biosynthesis. Compound **13** is then epimerized at C-3 and C-5, and finally converted to GDP-L-fucose (**16**) by the reduction of the C-4 ketone.

6-Deoxy-D-gulose is another unusual 6-deoxysugar which is a component of LPS from several different *Yersinia enterocolitica* strains. Based on sequence analysis, four genes are proposed to be involved in the biosynthesis of CDP-6-deoxy-D-gulose (**22**). The biosynthetic pathway is proposed to consist of the following reactions in sequence: activation of glucose-1-phosphate (**1**) to CDP-glucose (**19**), 4,6-dehydration to give CDP-

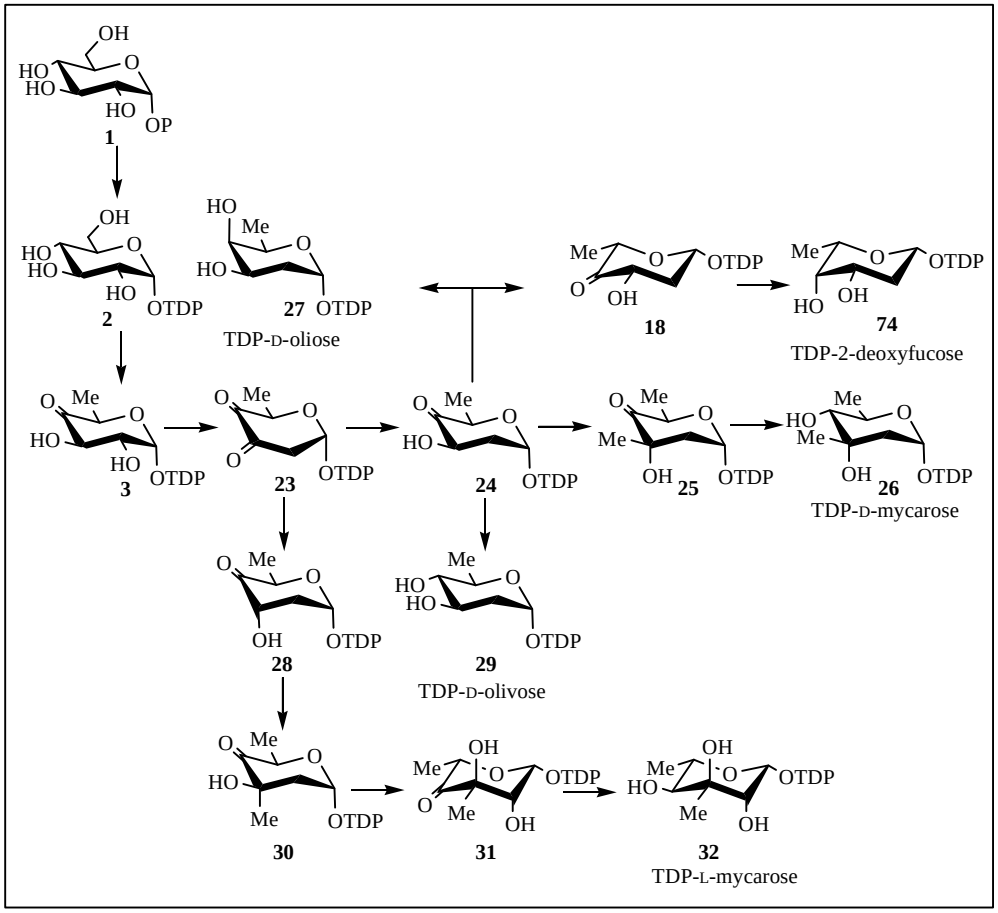
6-deoxy-4-keto-D-glucose (**20**), followed by 3-epimerization and 4-ketoreduction to give the final product.[24,95]

The aminocoumarin antibiotics coumermycin A_1, clorobiocin, and novobiocin, are produced by various *Streptomyces* strains and have demonstrated strong antibiotic potency against Gram-positive pathogenic bacteria. The three antibiotics have a common sugar unit, TDP-6-deoxy-5-methyl-L-*lyxo*-hexose (**7**). In the biosynthesis of this TDP-6-deoxysugar, the first two steps are glucose-1-phosphate (**1**) activation and TDP-glucose-4,6-dehydration, common to all TDP-deoxysugar biosynthetic pathways. Then the TDP-6-deoxy-4-keto-glucose (**3**) undergoes 3,5-epimerization (**4**), C-5 methylation (**6**) and 4-ketoreduction to form the final product, which is then transferred to the polyketide acceptor.[97,98]

Nogalose is a sugar unit attached to the polyketide portion of nogalamycin, an antitumor antibiotic produced by *Streptomyces nogalater*. TDP-6-deoxy-4-keto-glucose (**3**) is formed by the sequential catalysis of SnogJ and SnogK from glucose-1-phosphate (**1**). The next steps are a 3,5-epimerization (**4**) catalyzed by SnogF, C-3 methylation (**8**) catalyzed by SnogG2, and 4-ketoreduction (**9**) catalyzed by SnogC.[99]

D-Perosamine is often found as a component of the LPS of *E. coli*. However, it has also been discovered to be a component of the structures of some polyene antibiotics. In the biosynthetic pathway of GDP-D-perosamine (**17**), after the formation of GDP-6-deoxy-4-keto-mannose (**13**) by the identical route that has been described in the biosynthesis of GDP-L-fucose (**16**) and GDP-D-rhamnose (**14**), a transamination reaction at C-3 concludes the synthesis of GDP-D-perosamine (**17**).[24]

Biosynthetic Pathways of 2,6-Dideoxysugars (Scheme 1-2)



Scheme 1-2 . 2,6-Dideoxy hexose biosynthetic pathways.

2,6-dideoxysugars in which the hydroxyl groups at C-2 and C-6 have been replaced with either hydrogen or other non-oxygenated functional groups are often found as structural components of many secondary metabolites. As more biosynthetic gene clusters of deoxysugars are sequenced, it has been shown that the first two steps in the

deoxysugar biosynthesis are always hexose-1-phosphate activation to form an NDP-hexose and 4,6-dehydration to form NDP-6-deoxy-4-keto-hexose. It can also be generalized that the biosynthesis of 2,6-dideoxysugars and 3,6-dideoxysugars, 2,3,6-trideoxysugars and 3,4,6-trideoxysugars all have NDP-6-deoxy-4-keto-hexose as a common intermediate. From this point, the various deoxysugar biosynthetic pathways diverge.

Mithramycin, produced by *Streptomyces argillaceus*, has been used as an antitumor drug. As shown in Figure 1-6, the two sugar appendages that are attached to the aromatic polyketide of mithramycin include a disaccharide chain made up of two D-olivoses attached to a phenolic oxygen, and a trisaccharide chain made up of D-olivose, D-oliose, and D-mycarose attached to a β-hydroxyl group. The biosynthesis gene clusters of these three deoxysugars have been sequenced. Based on sequence analysis, the biosynthetic pathway of TDP-D-oliose (**27**) has been suggested. Following the formation of TDP-6-deoxy-4-keto-glucose (**3**), 2,3-dehydration and 3-ketoreduction lead to TDP-2,6-dideoxy-4-keto-glucose (**24**). A final 4-ketoreduction gives TDP-D-oliose (**27**). It should be noted that the hydroxyl group formed by the 3-ketoreduction is equatorial.[100] TDP-D-olivose differs from TDP-D-oliose only by the stereochemistry at C-4. The two deoxysugars share the same biosynthetic pathway up to the final ketoreduction step; the TDP-D-olivose (**29**) pathway has a 4-ketoreductase with the opposite stereospecificity than that of the TDP-D-oliose (**27**) biosynthetic pathway. D-Olivose is a component of many antibiotics, such as urdamycin from *Streptomyces fradiae*, simocyclinone from *Streptomyces antibioticus*, landomycin from *Streptomyces globisporus*, and others. The

formation of the 4-acetyl group on D-olivose in the structure of simocyclinone is proposed to take place after the sugar unit has been transferred to the aromatic polyketide.[100,101,102]

Both D-mycarose and L-mycarose have been found in antibiotic structures. D-

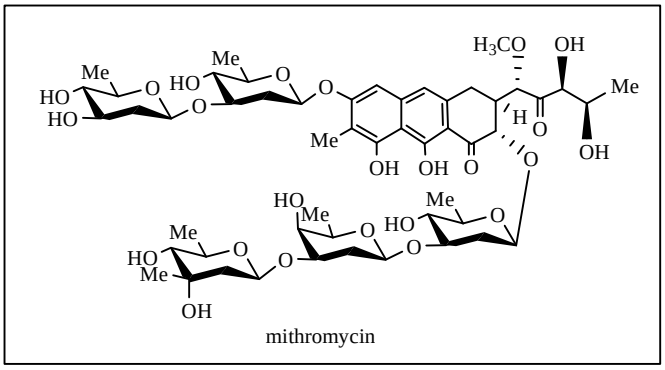


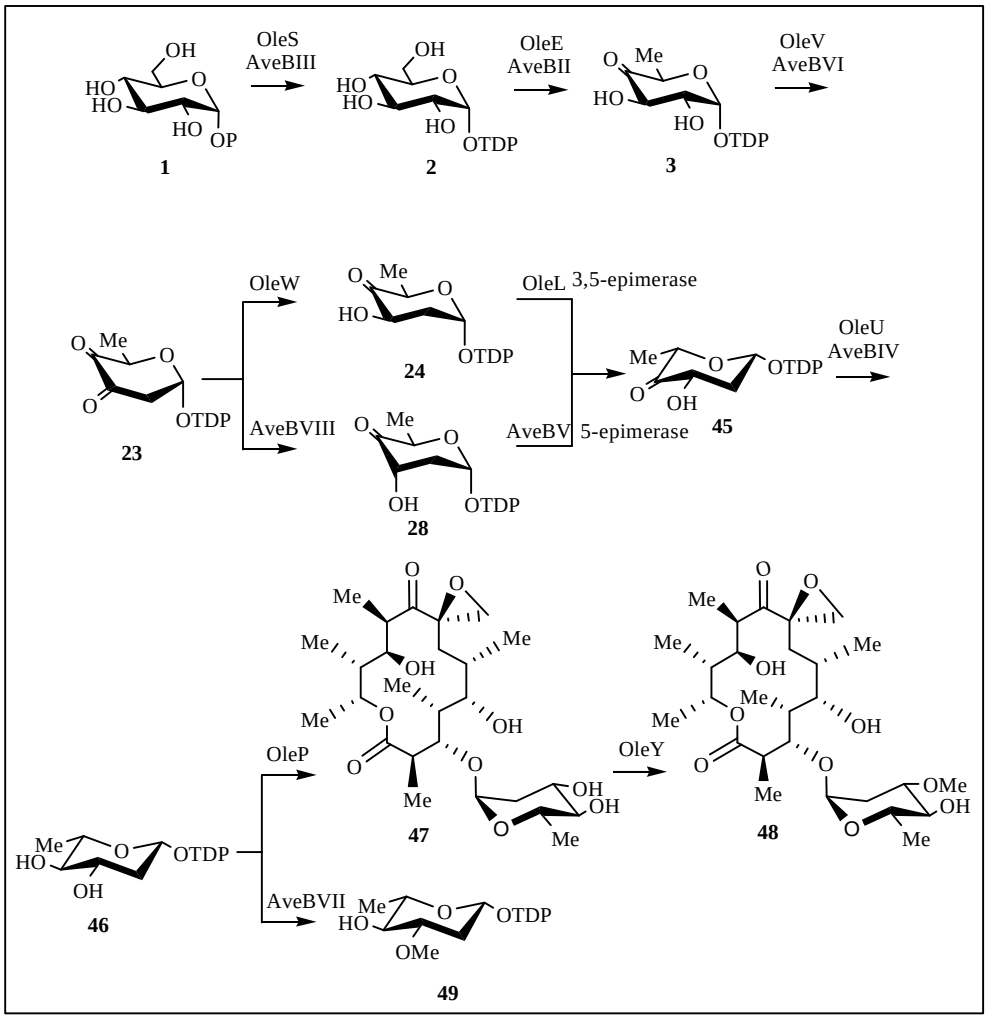
Figure 1-6. Structure of mithromycin

Mycarose is the terminal deoxysugar unit of the trisaccharide chain in mithramycin. It also occurs in chromocyclomycin which is an antitumor drug produced by various *Streptomyces* strains. The structure of chromocyclomycin is similar to that of mithramycin, having two oligosaccharide chains attached to an aromatic polyketide. In the biosynthesis of TDP-D-mycarose (**26**), two enzymes, 2,3-dehydratase and 3-ketoreductase, catalyze the C-2 deoxygenation on the intermediate TDP-6-deoxy-4-keto-glucose (**3**) to form TDP-2,6-dideoxy-4-keto-glucose (**24**) with an equatorial 3-hydroxyl group. Next, a methyl group is attached to C-3 of the hexose by the action of a *C*-methyltransferase. Finally, TDP-2,6-dideoxy-4-keto-3-*C*-methyl-hexose (**25**) undergoes 4-ketoreduction to form the final product TDP-D-mycarose (**26**). While the biosynthetic

pathway of D-mycarose is based on sequence analysis of the deoxysugar gene cluster, the biosynthesis of L-mycarose has recently been biochemically established due to the efforts of our group. From the common intermediate TDP-6-deoxy-4-keto-hexose (**3**), the 2-hydroxyl group is removed by the action of a 2,3-dehydratase and a 3-ketoreductase. The resulting 3-hydroxyl group in **28** is axial, unlike that of **24** in the biosynthesis of TDP-D-mycarose. Next, a methyltransferase catalyzes the attachment of a methyl to the C-3 position followed by 5-epimerization and 4-ketoreduction. The final product of TDP-L-mycarose has a different stereochemistry at C-3, C-4, and C-5 from that of TDP-D-mycarose. L-Mycarose has been found in many antibiotics, such as tylosin and erythromycin, which are produced by various *Streptomyces* strains.[103,104]

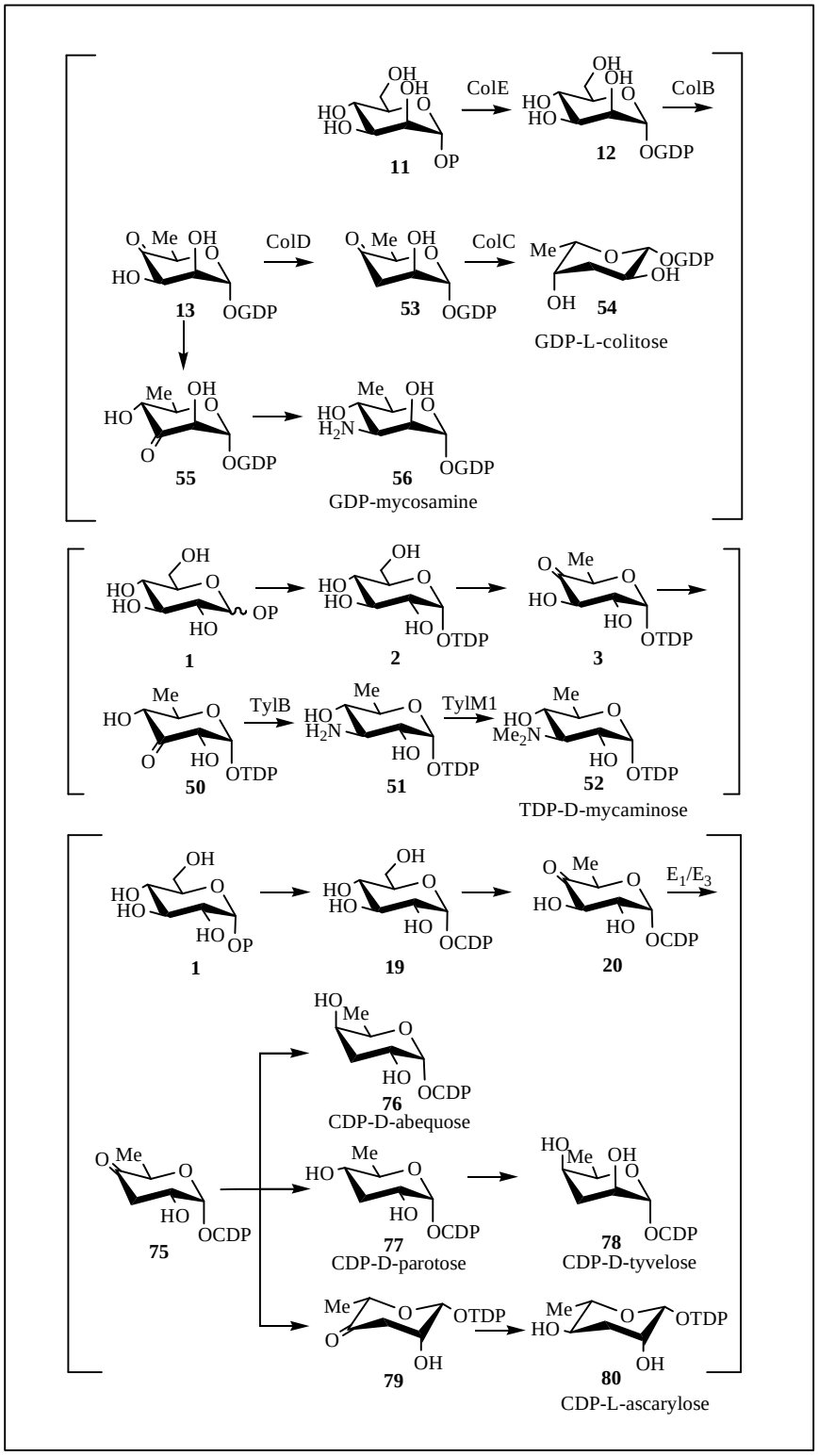
Another 2,6-dideoxysugar that often occurs in natural products is L-oleandrose. It has been found in the antibiotic oleandomycin produced by *Streptomyces antibioticus* and in the antiparasitic agent avermectin produced by *Streptomyces avermitilis*. The biosynthesis of TDP-L-oleandrose (**46**) is unique since it is the only known deoxysugar to be synthesized by two alternate routes in *Streptomyces antibioticus* and *Streptomyces avermitilis* (Scheme 1-3). Both routes begin with C-6 deoxygenation to form TDP-6-deoxy-4-keto-glucose (**3**), which then undergoes 2,3-dehydration catalyzed by OleV and AveBVI in the oleandomycin and avermectin pathway, respectively. Based on sequence analysis and genetic studies, C-3 ketoreduction catalyzed by OleW in the oleandomycin pathway gives an equatorial hydroxyl group in **24**, while that catalyzed by AveBVIII in the avermectin pathway gives an axial hydroxyl group in **28**. The same final product is

obtained by 3,5-epimerization catalyzed by OleL in the oleandomycin pathway and 5-epimerization catalyzed by AveBV in the avermectin pathway. Interestingly, evidence



Scheme 1-3. L-Oleandrose biosynthetic pathway.

from bioconversion experiments suggest that 3-*O*-methyltransfer occurs after glycosylation in the oleandomycin biosynthetic pathway, but before glycosylation in the avermectin biosynthetic pathway.[105,106,107]



Scheme 1- 4. 3,6-Dideoxy hexose biosynthetic pathways.

Biosynthetic Pathways of 3,6-Dideoxysugars (Scheme 1-4)

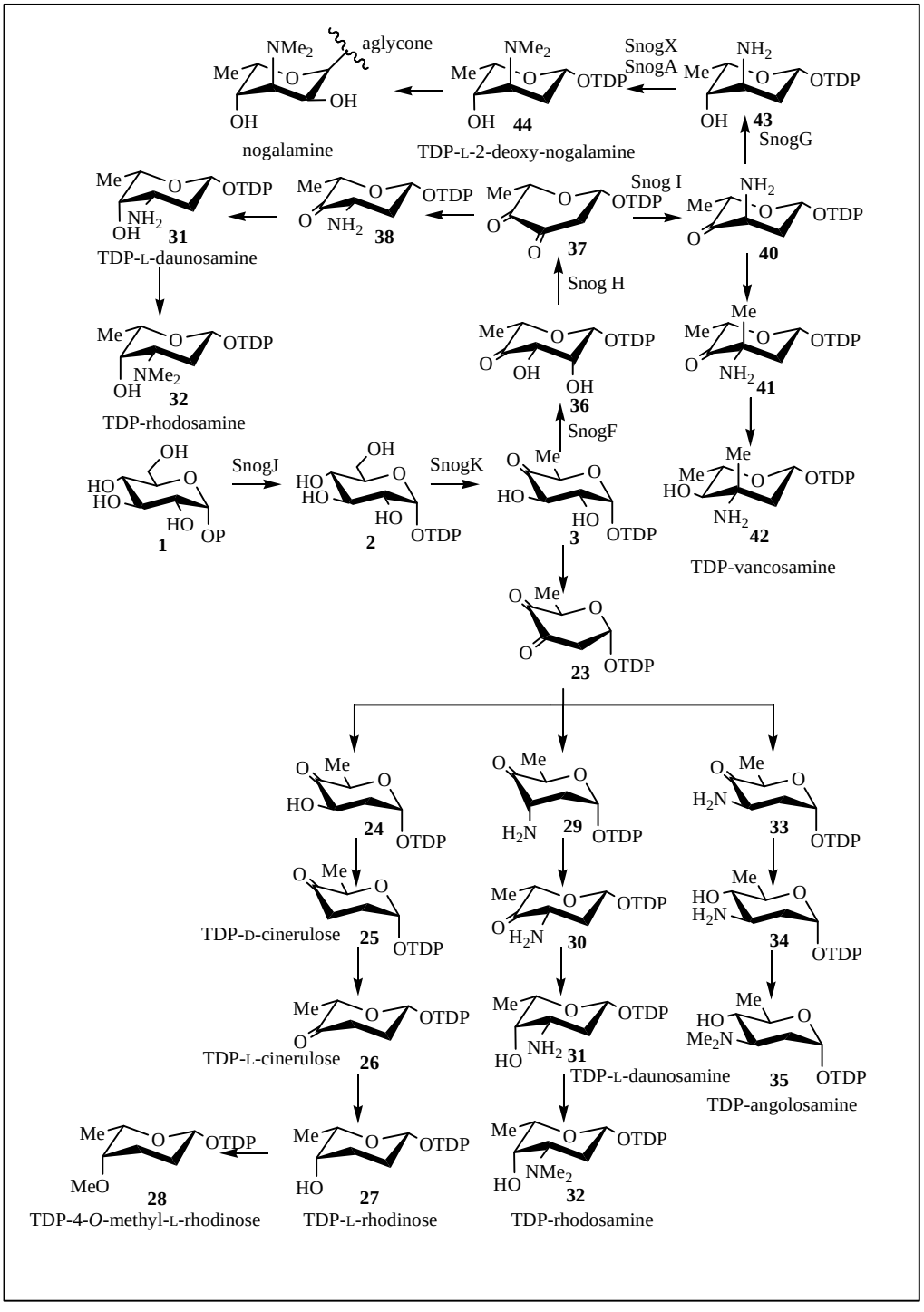
3,6-Dideoxysugars are widely found as components of LPS and surface layer glycoproteins,[24,108] and also occur in some secondary metabolites. According to the definition of deoxysugar, 3-amino-6-deoxy-hexoses are also included in this class. However, the enzymes and mechanisms implicated in C-3 amino transfer are quite different from those of C-3 deoxygenation. 3,6-Dideoxy-3-amino-hexoses are often found in antibiotic structures.

L-Ascarylose, D-abequose, D-paratose, and D-tyvelose are components of LPS from *Yersinia pseudotuberculosis* and different *Salmonella* strains. They have attracted much attention because of their highly immunogenic properties. The biosynthesis of these four 3,6-dideoxysugars begin with glucose-1-phosphate (**1**), which is activated to form CDP-D-glucose (**19**). After the formation of the common intermediate CDP-6-deoxy-4-keto-glucose (**20**), C-3 deoxygenation occurs, which is catalyzed by a combination of a 3-dehydrtase and a reductase. After C-3 deoxygenation, the four biosynthetic pathways diverge from each other: a 5-epimerization and a 4-ketoreduction lead to the formation of CDP-L-ascarylose (**80**); two ketoreductases with different stereospecificities catalyze the 4-ketoreduction of CDP-3,6-dideoxy-4-keto-glucose to form CDP-D-abequose (**76**) and CDP-D-paratose (**77**), respectively; and a 2-epimerase catalyzes the conversion from CDP-D-parotose (**77**) to CDP-D-tyvelose (**78**). It is worthwhile to note that although 3-, 5-, or 3,5-epimerization occurs often in the biosynthetic pathways of deoxysugars, as has been described for the biosynthesis of 6-deoxy and 2,6-dideoxysugars, the occurrence of 2- or 4-epimerases is relatively rare.[24,108]

L-Colitose is a component of the LPS of *E. coli* O111. The biosynthetic pathway of GDP-L-colitose (**54**) is quite similar to that of the CDP-L-ascarylose (**80**). The functional characterization of each enzyme in the GDP-L-colitose (**54**) biosynthetic pathway has been completed due to the efforts of our group.[69] First, GDP-D-mannose-4,6-dehydratase converts GDP-D-mannose (**12**) to GDP-6-deoxy-4-keto-mannose (**13**). Then C-3 is deoxygenated by the action of ColD. ColD is different from E_1 and its close homologues in that ColD does not require a reductase protein, like E_3, to catalyze this transformation. The final enzyme, ColC has dual functions as an epimerase and a reductase, catalyzing 5-epimerization and the subsequent 4-ketoreduction to form the product, GDP-L-colitose (**54**).[69]

D-Mycaminose has been found to be a part of the structure of the well-studied antibiotic tylosin.[86,109] This antibiotic is produced by *Streptomyces fradiae* and is commonly used as an antibiotic in veterinary medicine. In the biosynthesis of TDP-D-mycaminose (**52**), after the formation of TDP-6-deoxy-4-keto-glucose (**3**), an isomerase catalyzes the conversion of TDP-6-deoxy-4-keto-glucose (**3**) to TDP-6-deoxy-3-keto-glucose (**50**). Next, an amino group is attached to the C-3 position in a transamination reaction catalyzed by TylB. The last step is TylM1 catalyzed *N*,*N*-dimethylation.[109]

A polyene macrolide consists of a large macrolactone ring that has a characteristic series of conjugated double bonds, as well as an exocyclic carboxyl group and a deoxysugar. D-Mycosamine has been found to be the deoxysugar component in several antifungal polyene macrolides, such as pimaricin, nystatin A1, amphotericin B, and candicidin D, which are produced by different *Streptomyces* species. The proposed



Scheme 1- 5. 2,3,6- Trideoxy hexose biosynthetic pathways.

biosynthetic pathway of GDP-D-mycosamine (**56**) begins with the formation of GDP-6-deoxy-4-keto-mannose (**13**) in a manner identical to that previously described. Next, an isomerase whose encoding gene has not been identified, catalyzes the formation of GDP-6-deoxy-3-keto-mannose (**55**). Finally, an amino group is attached to the C-3 position of the hexose to form GDP-D-mycosamine.[110,111]

Biosynthetic Pathways of 2, 3, 6 -Trideoxysugars (Scheme 1-5)

The 2,3,6-trideoxysugars can be found in many antibiotics and antitumor drugs. The replacement of the three hydroxyl groups on the hexose with a range of functional groups creates a structurally diverse class of deoxysugars. Accordingly, natural products carry 2,3,6-trideoxysugars have shown to display interesting biological activities.[113]

The first example of a 2,3,6-trideoxysugar discussed here is the widely observed sugar L-rhodinose. L-Rhodinose is a component of many antibiotic and antitumor drugs, such as landomycin A, urdamycin A and granaticin.[70,112,113] Landomycin A, the principle secondary metabolite of *Streptomyces cyanogenus*, is the largest angucycline known thus far, and the most active landomycin derivative. As shown in Figure 1-7, L-rhodinose occurs as the third and the terminal sugar unit in the sugar chain attached to the polyketide of landomycin A. Urdamycin A is also an angucycline antibiotic, and is produced by *Streptomyces fradiae*. It consists of a polyketide moiety, an *O*-glycosidyl linked L-rhodinose moiety and a C-C connected trisaccharide chain. Granaticin and the related metabolites, dihydrogranaticin, granaticin B and dihydrogranaticin B are produced by *Streptomyces violaceoruber*. One D-olivose sugar unit is unusually attached

to the aromatic polyketide via two C-C bonds at 9-C and 10-C and an L-rhodinose is linked to the D-olivose by a conventional glycosidic bond.

In the proposed biosynthetic pathway of L-rhodinose, TDP-6-deoxy-4-keto-glucose (**3**) is formed first, and is then converted to TDP-2,6-dideoxy-4-keto-glucose (**24**) by the sequential catalysis of a 2,3-dehydratase and a 3-ketoreductase. Next, the 3-hydroxyl group is removed by a 3-deoxygenase to form TDP-2,3,6-trideoxy-4-keto-glucose (**25**), which is also called TDP-D-cinerulose. The final two steps in the biosynthetic pathway are 5-epimerization, the product of which is TDP-L-cinerulose (**26**), and 4-ketoreduction, which leads to the formation of TDP-L-rhodinose (**27**).

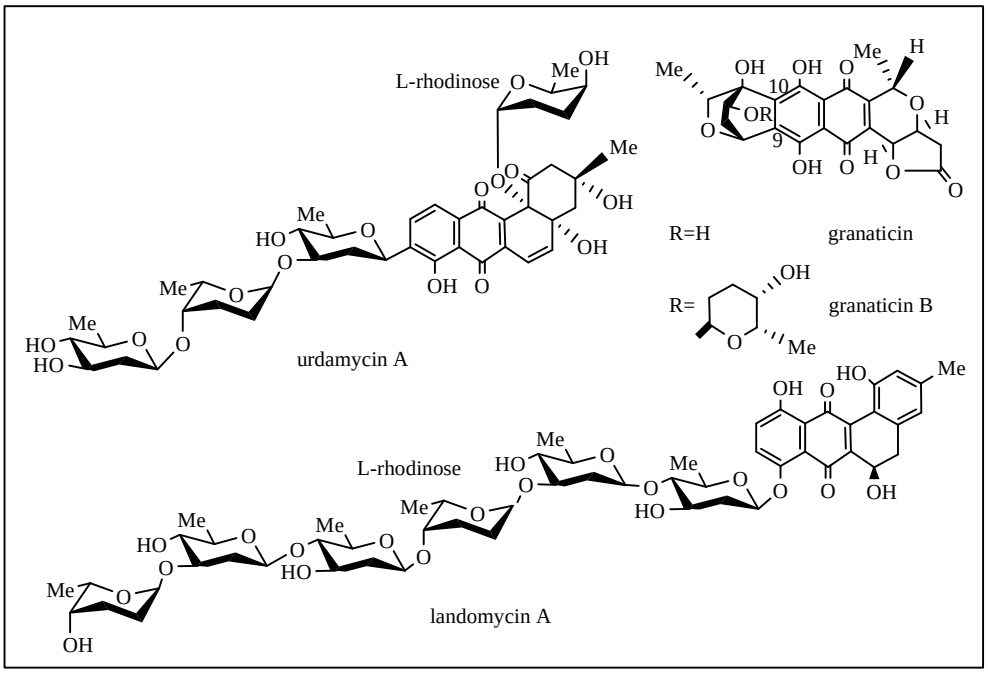


Figure1-7. Structures of urdamycin A, granaticin and landomycin A.

The polyether polyketide nanchangmycin is an antibiotic produced by *Streptomyces nanchangensis* NS3226. The complete biosynthesis gene cluster of the antibiotic has been sequenced.[114] 4-*O*-Methyl-L-rhodinose is a sugar component in the structure of nanchangmycin. The biosynthetic pathway of TDP-4-*O*-methyl-L-rhodinose (**28**) is proposed to be the same as that of TDP-L-rhodinose except an additional *O*-methylation step after the formation of TDP-L-rhodinose (**27**). From the preliminary analysis, the 4-hydroxyl group is methylated before the sugar unit is attached to the polyketide aglycone.[114]

Streptomyces peucetius and other *Streptomyces* strains are producers of daunorubicin and doxorubicin, both of which are clinically important drugs. Both of these two antibiotics contain the deoxysugar L-daunosamine. The daunorubicin biosynthetic gene cluster has been sequenced, and the functions of most of the genes in the cluster have been assigned based on sequence analysis. There have also been heterologous expression and bioconversion experiments showing the minimum set of genes required for daunosamine biosynthesis and attachment.[115] In the absence of any biochemical characterization of these gene products, two possible pathways for the biosynthesis of TDP-L-daunosamine (**31**) have been suggested. In both pathways, the activated TDP-6-deoxy-4-keto-glucose (**3**) is formed first. In one pathway, the steps following the formation of TDP-6-deoxy-4-keto-glucose are 3,5-epimerization (**36**), 2,3-dehydration (**37**), 3-transamination (**38**) and 4-ketoreduction (**31**), in that order. In the other proposed pathway, the reaction sequence is 2,3-dehydration to form TDP-6-deoxy-3,4-diketo-glucose (**23**), followed by direct 3-transamination (**29**), 5-epimerization (**30**)

and 4-ketoreduction (**31**). The two pathways differ in the step in which 5-epimerization occurs. The establishment of the correct biosynthetic pathway of TDP-L-daunosamine (31) must await biochemical characterization. However, it should be noted that in the deoxysugar biosynthetic pathways containing an epimerization step that have been fully or partially biochemically established, epimerization reactions generally occur in late steps.[24,115]

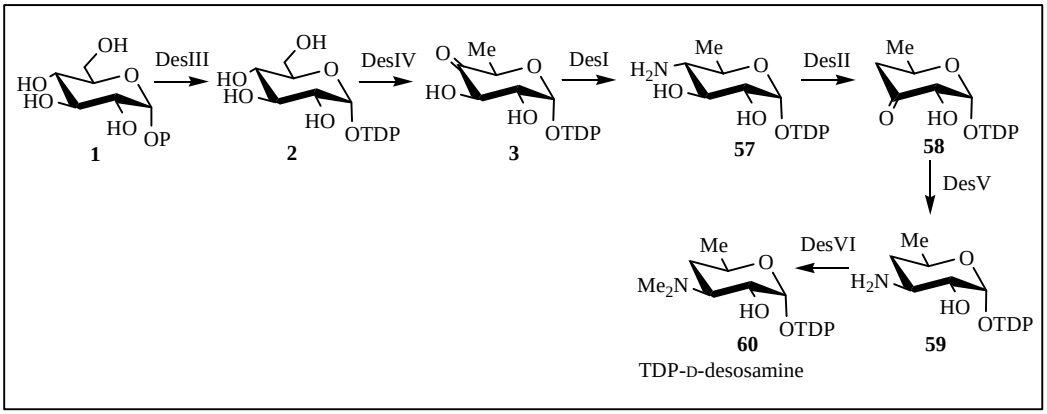
Aclarubicin, isolated from *Streptomyces galilaeus*, has remarkable cytotoxic activity and is a member of the anthracycline antibiotic family. Aclarubicin consists of an aklavinone aglycone and a trisaccharide moiety, which includes the 3-amino-2,3,6-trideoxysugar rhodosamine.[115] The structure of L-rhodosamine is similar to that of L-daunosamine but with two methyl groups attached to the amino. The methylation is suggested to occur before the sugar is transferred to the polyketide aglycone. Therefore, the biosynthetic pathway of TDP-L-rhodosamine (**32**) is proposed to be identical to that of TDP-L-daunosamine (**31**), but with the addition of a final *N*,*N*-dimethylation step.[115]

D-Angolosamine is the sugar component of the aromatic antibiotic medermycin. Medermycin was isolated from *Streptomyces sp*. AM-7161 and has proved to be very effective against gram-positive bacteria.[127] In the biosynthetic pathway of TDP-D-angolosamine (**35**), TDP-6-deoxy-4-keto-glucose (**3**) is proposed to undergo 2,3-dehydration (**23**) followed by 3-transamination (**33**). The amino group thus formed has an equatorial orientation (**33**). The subsequent 4-ketoreduction (**34**) and 3-*N*,*N*-dimethylation complete the formation of TDP-D-angolosamine (**35**).[127]

Nogalamycin is an anthracycline antibiotic produced by *Streptomyces nogalater*. Nogalamine is a 3,6-dideoxy-3-*N*,*N*-dimethyl-aminosugar that forms part of the structure of nogalamycin. The entire biosynthetic gene cluster of nogalamycin has been cloned and sequenced. The genes that are involved in nogalamine formation have been identified by gene deletion studies. Functions of some of the genes have been assigned based on the results from the deletion studies. Surprisingly, it was discovered that the sugar unit is synthesized as 2-deoxynogalamine before being transferred to the polyketide of nogalamycin.[116] The discovery of the genes involved in 2-deoxygenation in the sugar biosynthetic pathway is consistent with this conclusion. Thus, the proposed biosynthetic pathway of 2-deoxynogalamine has been suggested: the enzymes encoded by *snogJ* and *snogK* catalyze the formation of TDP-6-deoxy-4-keto-glucose (**3**) from glucose-1-phosphate (**1**) which then becomes the substrate of the 3,5-epimerase encoded by *snogF*. Next, a 2,3-dehydratase, probably encoded by gene *snogH*, catalyzes the formation of the unstable intermediate, TDP-2,6-dideoxy-3,4-diketo L-glucose (**37**) that undergoes transamination at C-3 by the action of the gene product of *snogJ*. The subsequent steps are 4-ketoreduction and 3-*N*,*N*-dimethylation. The enzyme encoded by *snogG* catalyzes the 4-ketoreduction based on sequence analysis while interestingly, both *snogA* and *snogX* encode putative *N*-methyltransferases. Thus, the attachment of the two methyl groups to the amino might be completed by two individual enzymes, which is different from the previously described *N*,*N*-dimethylations that are thought to be catalyzed by a single enzyme. The 2-oxygenase that is required to regenerate the 2-hydroxyl group at C-2 in nogalamine has not yet been identified.[116]

Vancomycin has been used as the last line of antibiotic to treat bacterial strains that have become drug-resistant. Vancomycin belongs to the glycopeptide antibiotic family whose aglycone is made of an oligopeptide with several crosslinked amino acids. A disaccharide chain is attached to the phenolic oxygen of the tyrosine residue of the aglycone. The biosynthetic gene cluster of vancomycin has been cloned and sequenced. A biosynthetic pathway for the terminal sugar of the disaccharide chain, named L-vancosamine, has been suggested. After the formation of TDP-6-deoxy-4-keto-glucose (**3**), 3,5-epimerization (**36**) followed by 2,3-dehydration leads to the formation of the unstable TDP-2,6-dideoxy-3,4-diketo-L-glucose (**37**), as was proposed for nogalamine biosynthesis. The keto group at C-3 then undergoes transamination (**40**) followed by 3-*C*-methylation (**41**) as occurs in nogalose, D- and L-mycarose biosynthesis. The final step is 4-ketoreduction to give TDP-L-vancosamine (**42**).[117]

Biosynthetic Pathways of 3,4,6-Trideoxysugars (Scheme 1-6)



Scheme 1-6. TDP-D-desosmaine biosynthetic pathway.

Desosamine has been found to be a sugar component of many natural products, such as erythromycin produced by *Saccharopolyspora erythraea*, oleandomycin produced by *Streptomyces antibioticus*, and methymycin, neomethymycin, narbomycin and picromycin produced by *Streptomyces venezuelae* (ATCC 15439). The presence of the desosamine is essential to the antibiotic potency and specificity of these compounds. Structural changes of this sugar unit in these compounds can severely compromise their capacities as antibiotics. The biosynthetic gene cluster of desosamine has been cloned and sequenced, and the biosynthetic pathway established due mainly to the efforts of our group. Genes *desIII* and *desIV* encode a TDP-D-glucose synthase and a TDP-D-glucose-4,6-dehydratase, respectively. These two enzymes together convert glucose-1-phosphate (**1**) to TDP-6-deoxy-4-keto-glucose (**3**). Next, the aminotransferase DesI catalyzes addition of an amino group to C-4. Subsequently, DesII catalyzes the formation of TDP-4,6-deoxy-3-keto-glucose (**58**). It had been unclear whether C-4 aminotransfer was a step in the biosynthetic pathway until *in vivo* gene deletion studies unambiguously established the functions of the gene products of *desI* and *desII*. Next, another aminotransferase encoded by *desV* catalyzes a transamination reaction at C-3. Finally, further 3-*N*,*N*-dimethylation leads to the formation of TDP-D-desosamine (**60**).[72]

Diverse sugar structures have been discovered during the past twenty years as a result of searching for more powerful antibiotics in order to counteract the increasingly serious problem of antibiotic resistance. These discoveries are also the result of intensive studies on cell-cell interaction, recognition and immunospecificity designed to find effective anti-tumor and anti-cancer drugs. Much progress has been made in these fields,

but much work remains. With ongoing research in relevant fields, increasing numbers of sugar structures are being identified. However, studies of the biosynthesis of naturally occurring sugars are in many cases inadequate, and conclusive information is limited. The number of the genes whose functions have been confirmed by *in vivo* gene deletion studies and/or by *in vitro* enzymatic characterization, is still small considering the vast amount of genes that might be involved in the biosynthesis of deoxysugars. Most of the biosynthetic pathways that have been discussed are only speculative, and are often solely based on sequence analysis. It is possible that the actual pathways may be different from those that have been suggested since some genes whose products are proposed to catalyze certain steps in the pathway, have not yet been verified. It is also possible that the reaction sequences that have been suggested may vary from the actual biosynthetic pathways.

3. C-O BOND CLEAVAGE IN THE BIOSYNTHESIS OF DEOXYSUGARS

As explained previously, deoxysugars, which have hydroxyl groups on the sugar replaced with hydrogens display different biological activities compared with their parent sugars, and are frequently utilized in the biosynthesis of bacterial cell wall and in the construction of natural products to convey immunogenic activity and biological specificity. Thus, the mechanistic logic involved in the C-O bond cleavage has attracted much attention and has been an active field of carbohydrate research.[52,53] Years of studies have expanded our understanding of the intricacies of deoxygenation mechanisms: no single mechanism can apply to all deoxygenation reactions; multiple

machineries are possible depending on the structure of the substrate and the site of deoxygenation on the sugar ring. A more underlying reason why different mechanisms are involved is that the carbon atoms where the hydroxyl groups are attached to are differentially activated. In the case of 2-deoxyribose biosynthesis, to break an unactivated C-O bond, a protein radical must previously be generated to facilitate the formation of the substrate radical.[76,77] As for C-O bond cleavage at C-2 and C-3 of a hexose, the substrate must be activated by the presence of a 4-keto functional group. C-2 and C-3 are at α and β positions relative to the keto group respectively. The removal of the hydroxyl group at C-2 usually goes through a β-elimination/reduction mechanism,[70,71] while that at C-3 needs further activation by the enzyme cofactor PMP/PLP.[63,64,69] The 4-keto group in the substrate first forms a Schiff base with the amine group of PMP. After deprotonation at C-4' of PMP, the high reduction potential of the PMP pyridine provides the driving force for the α-hydroxyl detachment. The overall reduction of the sugar substrate is balanced by the oxidation of another substrate. For the mechanisms that have been clarified, the other substrate can be either a simple L-glutamate which is oxidized to α-ketoglutarate, or it can be a reducing agent such as NAD(P)H, whose electrons are transferred through a reductase to the PMP-substrate complex by a radical mechanism. Although the substrate for C-O bond cleavage at C-6 position is not an activated keto-hexose, the well-studied mechanism has demonstrated that the first ministep in the reaction is indeed the formation of the 4-keto group which facilitates the later β-elimination/reduction. The 4-keto form is retained in the product structure, thus activating the sugar to achieve the subsequent conversions. The C-O bond cleavage at C-

4 is rare compared with that at other positions of the hexose. More evidence is required to completely decode the mechanism of 4-deoxygenation. Nevertheless, it has been well accepted that an activated PMP/PLP-substrate complex is an obligatory intermediate in its mechanism.[72]

C-O Bond Cleavage at the C-6 of Hexose (Figure 1-8)

C-6 deoxygenation, catalyzed by a NAD^+-dependent NDP-hexose 4,6-dehydratase, commonly named E_{od}, is the first committed step in the biosynthetic pathways of almost all deoxyhexoses. The NDP-hexose (**61**) is converted to the

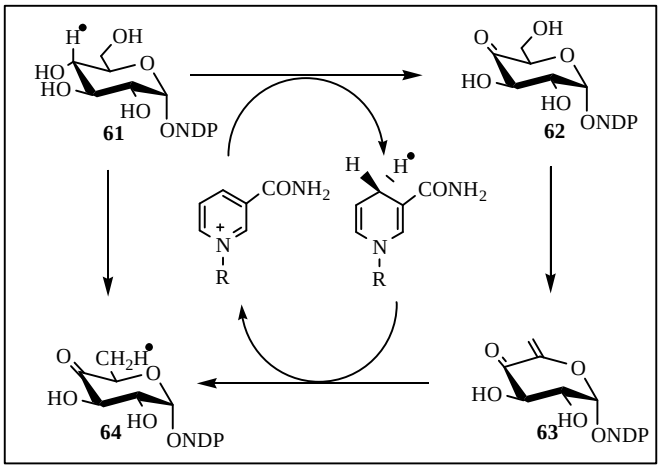


Figure 1-8 . Mechanism of C-6 deoxygenation catalyzed by E_{od}. Adapted from He, X. M. et al. Curr. Opin. Chem. Biol. 2002, 6, 590.

corresponding NDP-6-deoxy-4-keto-hexose (**64**) through the catalysis of E_{od}. The reaction involves a three-step sequential mechanism and requires the involvement of the cofactor NAD^+. The first step is the regiospecific oxidation of the substrate NDP-D-hexose to form NDP-4-keto-hexose (**62**) with the concomitant formation of NADH. A

water molecule is then eliminated across C-5 and C-6 (**63**). Since the first step has formed the activated NDP-4-keto-hexose, the water elimination in the second step is initiated by proton abstraction at C-5. The third step is the reduction of the NDP-4-keto-$\Delta^{5,6}$-glucoseen (**63**) intermediate by NADH to give the final product NDP-6-deoxy-4-keto-hexose (**64**). Clearly, E_{od} is a multifunctional enzyme, possessing keto oxidase, dehydratase and enoyl reductase activities.[52,53,118,119]

The analysis using stereospecifically radio-labeled NAD^+ has showed that the internal hydride transferred to and from NAD^+ is *pro-S* specific.[122] By monitoring the solvent isotope exchange, when the reaction is carried out in D_2O or $H_2{}^{18}O$, it has been concluded that the β-hydroxy elimination by dehydration is a concerted reaction.[123] Studies of the stereochemical course of this reaction have demonstrated that the dehydration across C-5/C-6 happens with *syn* elimination stereochemistry, while the reduction of the double bond in the enoyl intermediate proceeds in an *anti* manner.[120-122]

Because NAD^+ is regenerated at the end of each catalytic cycle, E_{od} belongs to a small group of enzymes whose pyridine nucleotide coenzyme is a genuine prosthetic group. Site-directed mutagenesis and kinetic studies have demonstrated the importance of residue Asn190, His232 and Tyr160 that are conserved between *E. coli* E_{od} and a mechanistically related enzyme, UDP-galactose-4-epimerase. Based on the available crystal structure of UDP-galactose-4-epimerase, it is speculated that the side chains of Asn190 and His232 may play important roles in facilitating hydride transfer between the sugar substrate and the nicotinamide cofactor; Tyr160 is likely to be the general base that deprotonates the 4-OH with the assistance of Thr134 and Lys164 via a proton relay

network. The importance of Asp135, Glu146, Glu198, Lys199 and Tyr301, which are only conserved among 4,6-dehydratases, have also been verified by mutagenesis and kinetic studies. [67,124-126,128]

C-O Bond Cleavage at the C-2 of Hexose (Figure 1-9)

The 2-hydroxyl elimination has been successfully characterized in the biosynthetic pathways of two deoxysugars, L-mycarose and L-rhodinose.[70,71] After TDP-6-deoxy-4-keto-glucose is formed, the hydroxyl group at C-2 is at the β position relative to the 4-keto group. As in the case of 6-deoxygenation, the detachment of the β-hydroxyl group at C-2 is achieved by eliminating a water molecule.

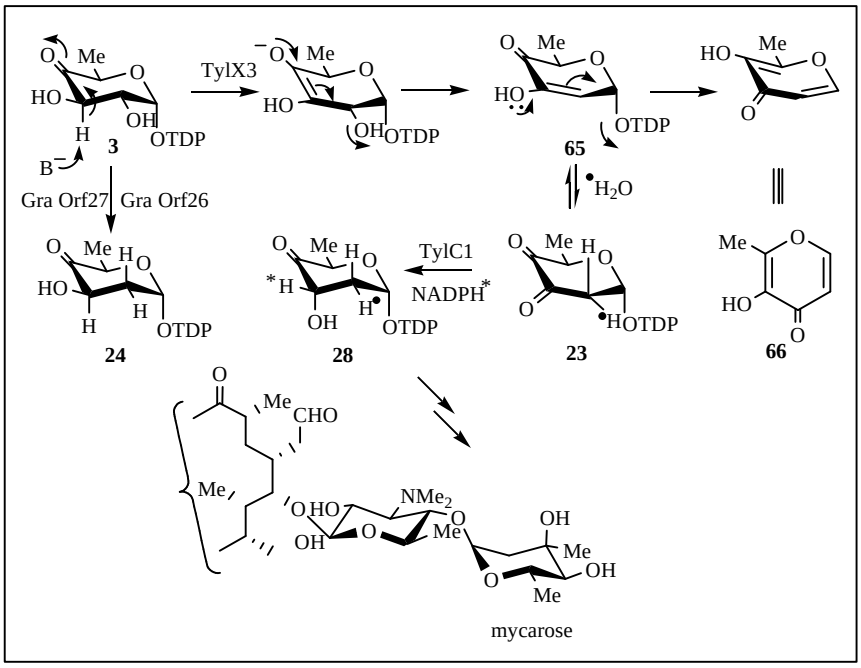


Figure 1-9 . Mechanism of C-2 deoxygenation. Adapted from He, X. M. et al. Curr. Opin. Chem. Biol. 2002, 6, 590.

For the biosynthesis of both L-mycarose and L-rhodinose, the α proton at C-3 position of TDP-6-deoxy-4-keto-glucose is acidic and can be easily abstracted by a base from the enzyme. The instability of the thus-formed enediolate intermediate drives the elimination of the 2-hydroxyl group to occur in either a concerted or a stepwise manner (**65**). The enzymes catalyzing the dehydration step are TylX3 and Gra Orf27 in the biosynthetic pathways of L-mycarose and L-rhodinose, respectively. Some initial characterization of TylX3 has shown that the enzyme is a Zn^{2+}-dependent dehydratase.[71] The Zn^{2+} ion could be responsible for polarizing the C-4 keto group of the substrate in order to ease the generation of an enolate, which would be used to expel the 2-hydroxyl group. Another possibility is that it could serve as an active site Lewis acid to assist the departure of the hydroxyl group at C-2. The actual role of Zn^{2+} in TylX3 catalysis is still not understood and further investigation is required. TylX3 has a high degree of sequence similarity to several deduced enzymes from antibiotic biosynthetic gene clusters, such as EryBVI in the erythromycin pathway, LanS in the landomycin pathway, Orf23 in the vancomycin pathway, SnogH in the nogalamycin pathway, and DnmT in the daunorubicin pathway. Accordingly, it is possible that each of these enzymes plays a role as a Zn^{2+}-dependent dehydratase catalyzing C-2 deoxygenation in each respective deoxysugar biosynthetic pathway.[47,99,103,113,117,116]

The dehydration product of TylX3 (**65**) is unstable. Without the immediate action of the following enzyme in the biosynthetic pathway, the dehydration product easily degrades to form maltol (**66**) and TDP.[70,71] The identity of maltol (**66**) has been confirmed by comparing the HPLC retention time and ^{1}H-NMR to those of the authentic

sample. Therefore, the following reduction step must happen instantly after the dehydration product is formed. In the biosynthesis of L-mycarose and L-rhodinose, TylC1 and Gra Orf26 are the reductases in their respective pathways. It has been assumed that TylX3 and TylC1 form a complex *in vivo* and that the dehydration product (**65**) formed by the catalysis of TylX3 is transferred to the reductase TylC1 through an internal channel. However, yeast two-hybrid analysis failed to detect an interaction between TylX3 and TylC1. This might be because the interaction between the two proteins is fairly weak, and that the two-hybrid system is not sensitive enough to detect the interaction. The existence of a protein–protein interaction between the dehydratase TylX3 and the reductase TylC1 is a favorable explanation to the puzzle of how degradation of the dehydration product is avoided in the 2-deoxysugar biosynthesis. More examination of the interaction between the two proteins is necessary. However, in the absence of solid evidence for the interaction, other explanations should not be excluded.

Although 2-deoxygenation in the biosynthesis of L-mycarsoe and L-rhodinose occurs through the same β-elimination/reduction mechanism, the stereochemistry of the reformed 3-hydroxyl group in the product is different. The hydroxyl group in the Gra Orf27/Gra Orf26 product (24) is equatorial while that in the TylX3/TylC1 product (28) is axial. Having an intermediate with a relatively high-energy axial group in the L-mycarose pathway could be beneficial in that it may facilitate the subsequent C-3 methylation reaction which is not included in the biosynthesis of L-rhodinose.

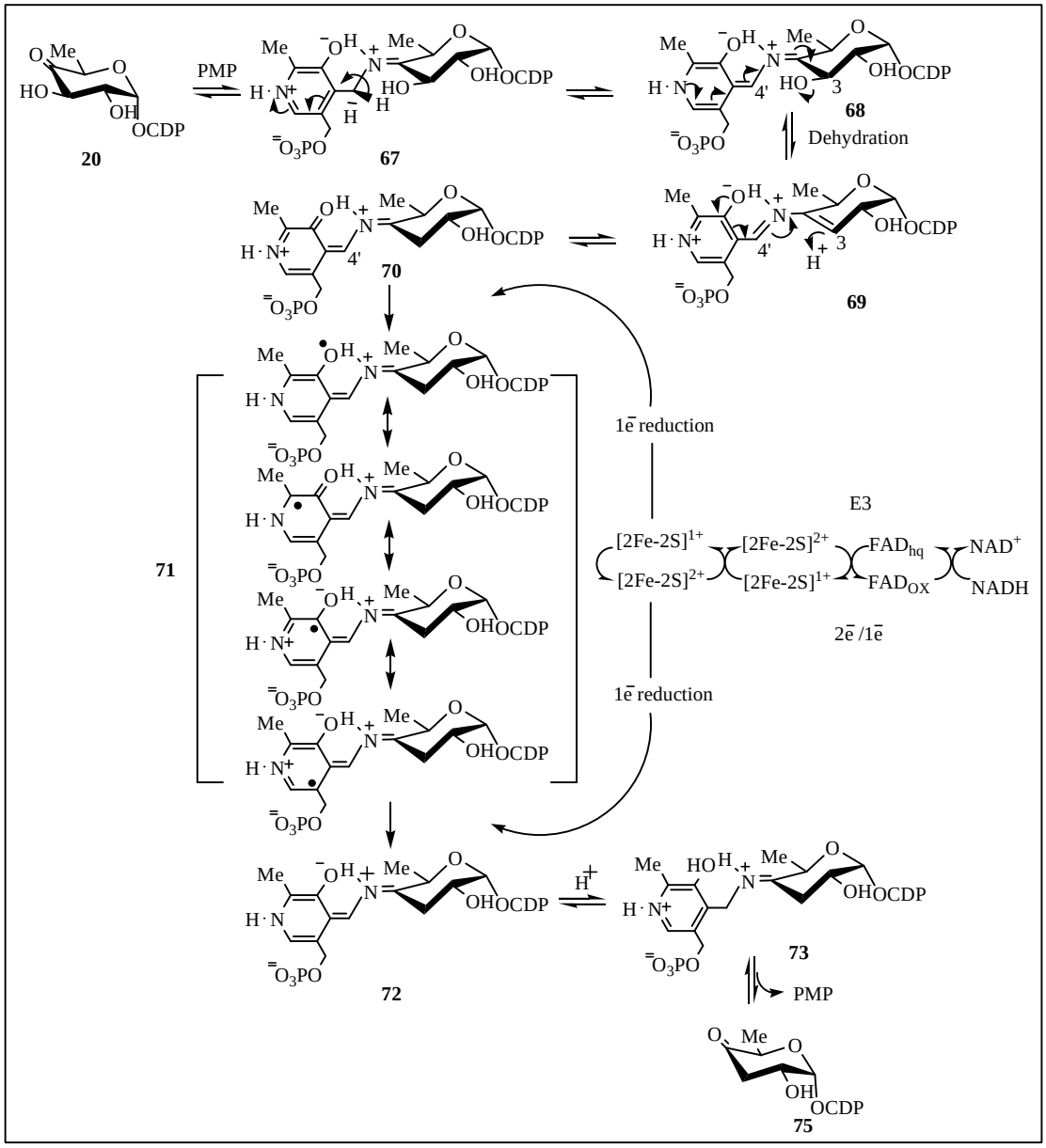


Figure 1-10 . E_1/E_3 catalyzed C-3 deoxygenation. Adapted from He, X. M. et al. Curr. Opin. Chem. Biol. 2002, 6, 590.

TylC1 acts as a typical ketoreductase and contains a Rossman fold (^{277}GXXGXXXG^{284}) near its C-terminus. It is likely a member of the long-chain alcohol-

dehydrogenase (ADH) family. The TylC1 reaction with (4*S*)-[4'-^{2}H]NADPH as reductant has indicated that the hydride transfer from NADPH is *pro-R* specific.[129] While TylX3 and Gra Orf27 share substantial sequence similarity, TylC1 and Gra Orf26 lack significant sequence similarity, which could explain the different stereochemistry of the hydroxyl groups in their products and might also indicate that the two enzymes have different evolutionary origins.

C-O Bond Cleavage at the C-3 of Hexose

E1/E3 Radical Mechanism (Figure 1-10)

The biosynthesis of 3,6-dideoxy hexoses has drawn particular attention. 3,6-Dideoxyhexoses have been found primarily at the non-reducing terminus of the *O*-antigen of the lipopolysaccharide on the outer cell membrane of many Gram-negative bacteria. The presence of 3,6-dideoxyhexoses in the *O*-antigen is considered to be the primary determinant of microbial pathogenicity and virulence. The 3,6-dideoxyhexoses, like 2,6-dideoxyhexoses, are derived from the common precursor NDP-6-deoxy-4-keto-hexose. However, deoxygenation at C-3 is mechanistically different from those at C-2 and C-6. The deoxygenation at C-2 and C-6 eliminates the β-OH group relative to the 4-keto in the hexose. The 4-keto group is either present in the substrate as in the case of C-2 deoxygenation, or is formed in the earlier catalytic steps as in the case of C-6 deoxygenation. In contrast, the C-3 deoxygenation occurs at the α-carbon relative to the 4-keto group. Because of the efforts of our group, along with the early work of Strominger and coworkers, the enzymatic mechanism of the C-3 deoxygenation in the

biosynthesis of CDP-ascarylose has been unraveled. The 3-deoxygenation on CDP-6-deoxy-4-keto-hexose is mediated by two enzymes: CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydratase (E_1) and CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydratase reductase (E_3). The E_1 enzyme is a pyridoxamine 5'-phosphate (PMP) dependent, [2Fe-2S] containing dimeric enzyme. The [2Fe-2S] cluster exhibits slightly rhombic symmetry (g_1 = 2. 012; g_2 =1.950; and g_3 = 1.932).[58,65] Different from other coenzyme B_6-dependent enzymes, whose cofactor pyridoxal 5'-phosphate (PLP) is covalently linked to a conserved lysine residue, PMP in E_1 is non-covalently bound and the conserved lysine residue is replaced with His220.[55] In the E_1 catalyzed reaction, PMP first forms a Schiff base with the 4-keto group of the substrate (67). Then, the *pro-S* C-4' hydrogen of the PMP-ketimine complex is abstracted by base His220 (68), triggering the expulsion of the C-3 hydroxyl group and producing the transient PMP-$\Delta^{3,4}$-glucoseen intermediate (69).[63] The nearby His222 may substitute for His220 as base, since the H220A mutant of E_1 still demonstrates certain activity.[55] When the reaction was conducted with radio labeled PMP, (4*S*)-[4'-^{3}H]PMP or (4*R*)-[4'-^{3}H]PMP, only the solvent in the reaction with (4*S*)-[4'-^{3}H]PMP was enriched with the radioactivity. This demonstrates that the proton abstraction from the C-4' of the PMP-ketimine complex (67) is *pro-S* specific.[63] It was also found that the replacement of the 3-OH group with a solvent hydrogen proceeds with retention of the configuration.[63] Up to this point, the E_1-catalyzed reaction is reversible. Without further driving force, the desired product CDP-3,6-dideoxy-4-keto-glucose could not be stably formed. The reversibility of the E_1 catalyzed reaction was confirmed by the reaction conducted in [^{18}O]H_2O.[63] Both the C-3

and C-4 positions of the recovered substrate incorporated ^{18}O as judged by GC-MS, with the equilibrium strongly favoring the starting material. The reduction of the glucoseen intermediate followed by the release of the 3,6-dideoxysugar is achieved by the other enzyme, E_3.

E_3 is an electron transfer protein, closely resembling many iron-sulfur flavoproteins in the ferredoxin-$NADP^+$ reductase (FNR) family.[56] The reduction of the glucoseen intermediate is most effectively mediated by E_3, but other electron transfer systems such as diaphorase and methane monooxygenase are competent as well, but with much lower efficiency.[63] Chemical reductant dithionite can also provide the required electrons and transfer them directly to the glucoseen intermediate under anaerobic conditions.[66] The CDP-3,6-dideoxy-$\Delta^{3,4}$-glucoseen intermediate (69) is reduced by the two electron equivalents in a step-wise manner. The ultimate reducing agent NADH provides the two electron equivalents in a hydride form to the FAD cofactor of E_3. The cofactor FAD, which is capable of transferring either a single electron or two electrons, functions as the $2e^-/1e^-$ switch and delivers the electrons one at a time to the [2Fe-2S] center, which is an obligatory one-electron redox coenzyme in E_3. The single electron is next delivered to the [2Fe-2S] center in E_1, where the electron is used to reduce the glucoseen intermediate. Since the complete reduction of the glucoseen intermediate requires two electron equivalents, the molar ratio between the substrate CDP-4-keto-6-deoxy-D-glucose and NADH should be 1:1. The E_1/E_3-catalyzed reaction involves single electron transfer and entails the formation of radical intermediates (71). Characterization of the assumed radical intermediates has been pursued via stopped-flow

spectrophotometry and rapid freeze-quench EPR spectroscopy.[62,68] When [4',5'-$^2H_4$]PMP-reconstituted E_1 was used in the reaction, the radical signal was narrowed by 3 G as compared with that of a reference spectrum obtained by using non-labeled PMP. Moderately large ^2H hyperfine couplings were found to exist between the unpaired radical and the deuterium of the labeled PMP in pulsed ENDOR spectroscopy. The combination of these studies provides compelling evidence for the radical formation in the mechanism of the E_1-E_3 reaction and unambiguously locates the radical on the PMP skeleton.

E_1 shows a high degree sequence similarity to quite a few enzymes that are suggested to be involved in the biosynthesis of 3,6-dideoxysugars. The elucidation of the E_1/E_3 mechanism may be useful in understanding the functions of homologous enzymes and may be beneficial in establishing other 3,6-dideoxysugar pathways. Mechanistically, the involvement of PMP in a radical-mediated reduction makes E_1 unique among coenzyme B_6-dependent enzymes.

α-Elimination by Coenzyme B_6 (Figure 1-11)

Several genes in the 3-deoxysugar biosynthesis gene clusters are considered to encode 3-deoxygenases based on their sequence homology to the sequence of E_1. The

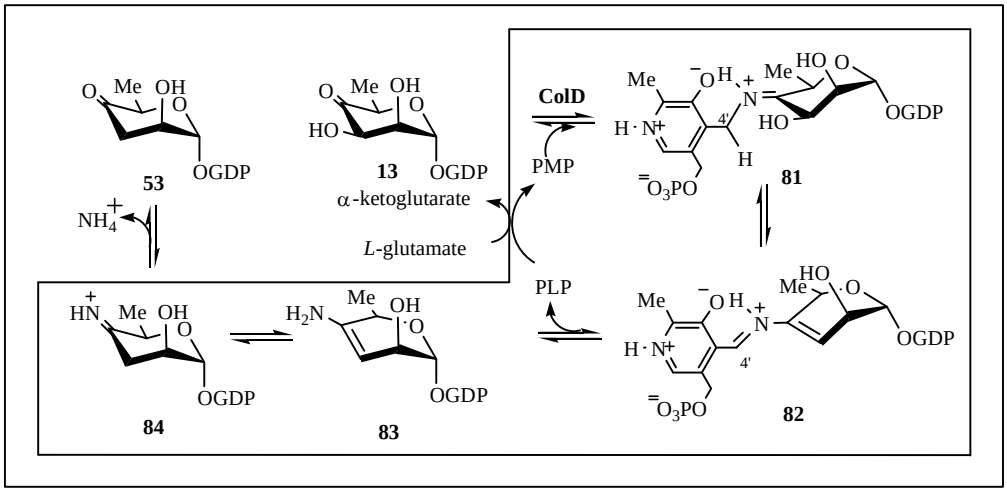


Figure 1-11 . ColD catalyzed C-3 deoxygenation. Adapted from Beyer, N. et al. J. Am. Chem. Soc. 2003, 125, 5584

enzyme encoded by *colD* in the biosynthetic gene cluster of GDP-L-colitose, is also assigned as a 3-deoxygenase, but the sequence similarity among ColD and E_1 and its homologs is only moderate. The fact that ColD alone catalyzes 3-deoxygenation might suggest that the ColD-catalyzed reaction in the biosynthetic pathway of GDP-L-colitose, a 3,6-dideoxyhexose found in the *O*-antigen of the gram-negative lipopolysaccharides, has a different mechanism from the E_1-catalyzed reaction. ColD is a coenzyme B_6-dependent enzyme. However, the lysine residue that is conserved among PLP-dependent enzymes and is the site of linkage to PLP, is replaced by a histidine residue in ColD. Such a substitution also occurs in E_1, and has been considered a characteristic unique to PMP-dependent enzymes. Different from the E_1/E_3 catalyzed reaction, the ColD-catalyzed reaction requires only an amino donor and no auxiliary protein. These differences are reasonable in light of the low level of sequence identity between ColD

and E_1, and are also consistent with the fact that a gene encoding E_3 homolog was not identified from the L-colitose biosynthetic gene cluster.[69] In the first step of the ColD reaction mechanism, the cofactor PMP is thought to form a Schiff base (81) with the 4-keto group of GDP-6-deoxy-4-keto-mannose (13). Next, deprotonation of C-4' of the ketimine complex forces the repulsion of the 3-hydroxyl group (82), which is at the α position relative to the keto group. Next, PLP is released from the complex and regenerated to become PMP, with the concomitant conversion of L-glutamate to α-ketoglutarate. Finally, by the action of water, the ammonia is discharged from the hexose (84) and the 4-keto group is restored. The overall reaction can be considered as a α-elimination reaction, one of the various types of reactions in which the versatile coenzyme B_6 plays an important role. Further experiments showed that L-glutamate is a better amino donor than L-glutamine and L-alanine, while D-glutamate and L-aspartate are inactive as cosubstrates.[130] The reversibility and stereochemistry of the reaction was studied by incubation with radioactivity-labeled PMP.[131] When the reaction was conducted with (4*R*)-[4'-^{3}H]PMP, radioactivity could be found in the solvent after the substrate, cofactor, and enzymes were removed. Unlike E_1, the *pro-R* hydrogen at C-4' of the ketimine is removed in the reaction. The utilization of PMP instead of PLP in a dehydration reaction is rare. Also, its combined deoxygenation-transamination activity distinguishes ColD as a unique enzyme.[69]

C-O Bond Cleavage at the C-4 of Hexose (Figure 1-12)

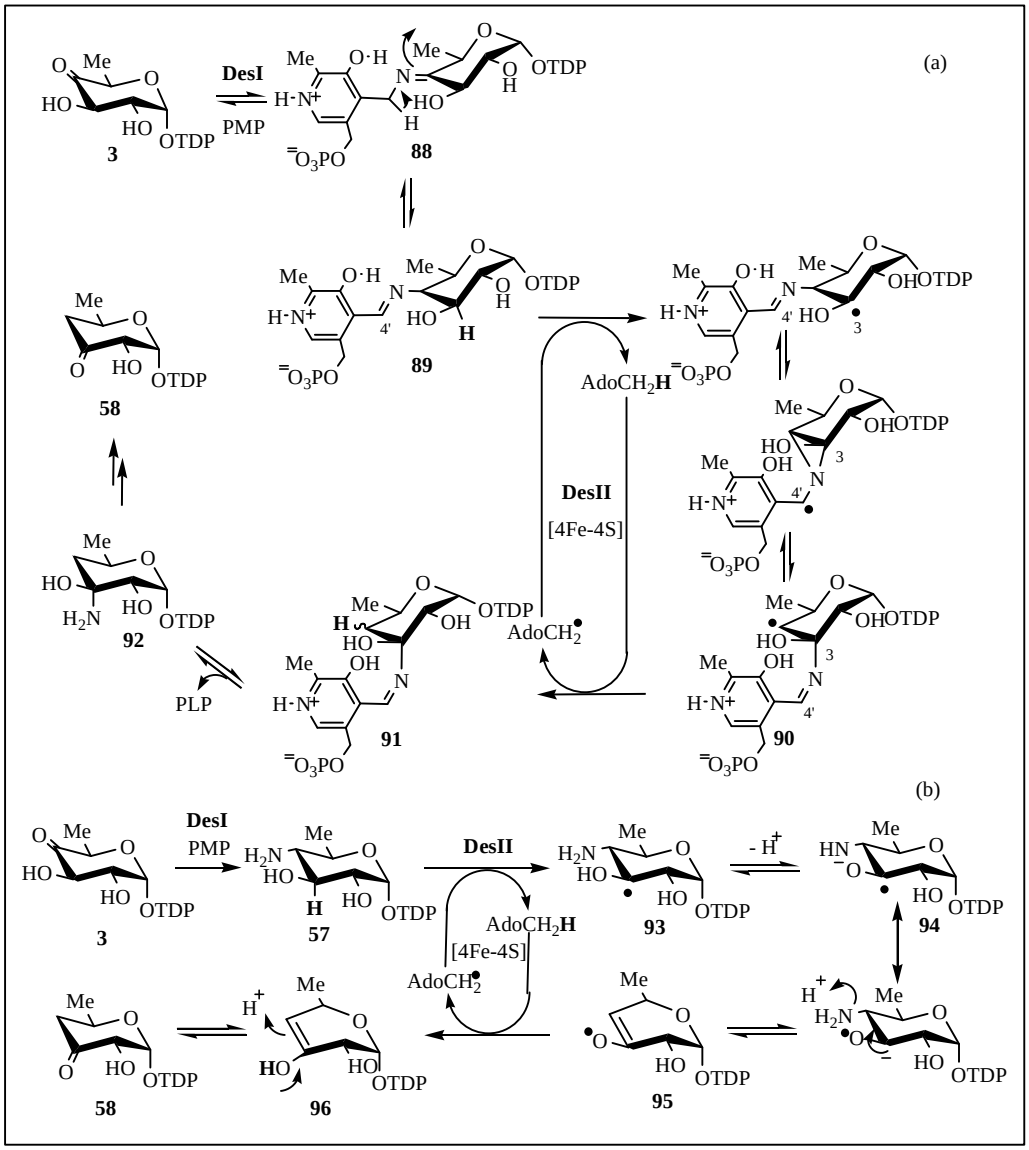


Figure 1-12 . DesII (DesI) catalyzed C-4 deoxygenation. (a) collaboration between DesI and DesII. (b) sequential action of DesI and Des II. Adapted from He, X. M. et al. Curr. Opin. Chem. Biol. 2002, 6, 590.

The conversion of TDP-6-deoxy-4-keto-glucose (3) to TDP -4,6-dideoxy-3-keto-glucose (58) is an intriguing process in the biosynthetic pathway of desosamine, a sugar component of several antibiotics produced by various *Streptomyces* strains, such as methymycin, pikromycin, and erythromycin. The TDP-desosamine biosynthetic gene

cluster has been cloned and sequenced. Based on sequence analysis, the *desIII* encoded enzyme is a glucose-1-phosphate thymidylyltransferase while that encoded by *desIV* is a TDP-glucose-4,6-dehydratase. The enzymatic functions deduced from the sequences of *desV*, *desVI* and *desVII* are aminotransferase, methyltransferse, and glycosyltransferase, respectively. They are the enzymes catalyzing the late steps in the pathway. From sequence analysis, the enzyme encoded by *desI* has a coenzyme B_6 binding site and that encoded by *desII* should contain an iron-sulfur center. Initially, it was proposed that 4-hydroxyl group elimination begins with isomerization, possibly catalyzed by the enzyme encoded by *desVIII*. The isomerization product, TDP-6-deoxy-3-keto-glucose would then undergo α-hydroxyl group elimination to remove the 4-OH group which is at the α-carbon relative to the 3-keto group. It was assumed that the deoxygenation mechanism should be similar to that of the E_1/E_3 reaction: DesI would play the role of dehydratase by forming a PMP-substrate complex, while DesII would act as a reductase analogous to E_3, transferring electrons from a reducing agent, possibly NADH. However, neither the proposed dehydrase/reductase function of DesI/DesII nor the isomerase function of DesVIII could be demonstrated by *in vitro* enzymatic assays. In *vivo* gene disruption experiments were carried out for the genes encoding DesI and DesII and yielded interesting results. If C-4 deoxygenation followed a similar path to C-3 deoxygenation, the phenotypes of *desI* and *desII* deletion mutants should have been the same. However, while macrolide analogues carrying a quinovose (85) were isolated from the fermentation broth of the K*desI* mutant, macrolides containing an *N*-acetylated 4-aminosugar (**86**)

were produced by the K*desII* mutant (Figure 1-13).[49,72] The fact that deletion of *desII* leads to the

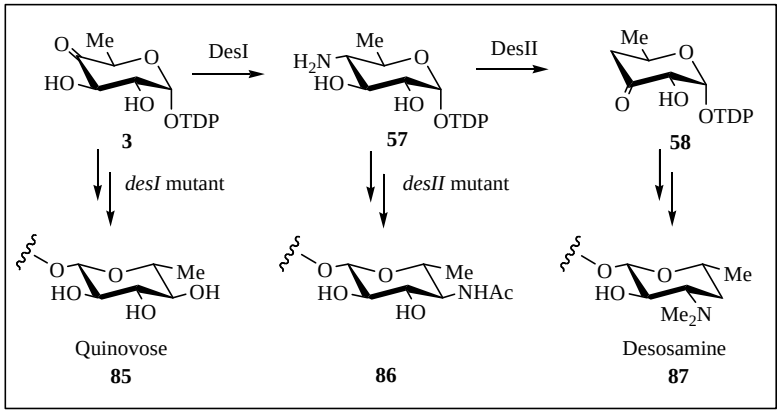


Figure 1-13 . *In vivo des*I and *des*II deletion studies. Adapted from He, X. M. et al. Curr. Opin. Chem. Biol. 2002, 6, 590.

production of a 4-aminosugar provided initial evidence indicating that DesI is a 4-aminotransferase. The *in vitro* enzymatic reaction, including the E_{od} product, DesI and an amino donor, such as L-glutamate, indeed produced TDP-4-amino-4,6-dideoxy-glucose. The formation of an amino group at the site where the deoxygenation should occur differentiates the mechanism of C-4 deoxygenation by DesI/DesII from that of C-3 deoxygenation catalyzed by E_1/E_3.

Bioinformatic analysis by H. J. Sofia lists DesII as a SAM radical-dependent enzyme,[132] although the consensus peptide sequence where SAM binds has not been ascertained for DesII. However, X-ray crystallographic studies revealed that SAM radical binds close to a [4Fe-4S] center in biotin synthase, another SAM radical enzyme which has been well studied. Most SAM radical dependent enzymes contain a [4Fe-4S]

center, the binding motif of which has been established as CXXXCXXC. In the sequence of DesII, this [4Fe-4S] binding motif has been identified. This provides additional evidence that DesII is not a homolog of E_3, which contains a [2Fe-2S] center.

In summary, C-4 deoxygenation may require SAM and the [4Fe-4S] center from DesII and the coenzyme B_6 from DesI, and go through a radical mechanism. Although the detailed mechanism is still not clear, there are two propositions: one is that DesI and DesII collaborate in a single step (Figure 1-12 (a)); the other is that each enzyme catalyzes a separate step and the DesI product is the substrate of the DesII catalyzed reaction (Figure 1-12 (b)). There has been one example where cofactors PLP, SAM, and [4Fe-4S] are all implicated in one reaction: the well-studied lysine 2,3-aminomutase catalyzed interconversion between L-lysine and L-β-lysine.[133,134] DesI and DesII may cooperate to have a reaction mechanism similar to this reaction: the cofactor PMP from DesI may form a Schiff base (88) with the 4-keto group in TDP-6-deoxy-4-keto-glucose. The 5'-adenosyl radical then abstracts the hydrogen at C-3 through homolytic cleavage of the C-H bond and generates a radical at C-3 of the sugar. By rearrangement, the radical first moves to the C-4' of PMP and then to the C-4 of the sugar (90), while at the same time, the linkage site between PMP and the hexose is switched from the C-4 to the C-3 of the hexose. The radical at the C-4 of the sugar then abstracts the hydrogen that has been taken away by the 5'-adenosyl radical to regenerate the SAM radical and complete the C-O bond cleavage at C-4 (91). In subsequent steps, PLP is released from the ketimine complex and elimination of a molecule of ammonia yields the final product TDP-4,6-dideoxy-3-keto-glucose (58). In the alternate hypothesis DesI and DesII work

independently. DesI is a 4-aminotransferase catalyzing the formation of TDP-6-dideoxy-4-amino-glucose (57), while DesII catalyzes the elimination of the 4-amino group. The DesII-bound SAM undergoes reductive cleavage to form the 5'-adenosyl radical, which initiates the reaction. The 5'-adenosyl radical abstracts the hydrogen at C-3 of the hexose through homolytic cleavage of the C-H bond to generate a radical at the C-3 of the sugar (93). Next, the hydroxyl group at C-3 is deprotonated to form an anion-radical intermediate (94). Electron rearrangement converts the oxygen anion-carbon radical to the carbon anion-oxygen radical. The instability of the latter drives the elimination of the amino group at the C-4 of the sugar. Subsequently, the oxygen radical (95) abstracts the hydrogen that has been taken by the 5'-adenosyl radical. The final rearrangement from the enolate (96) to the keto form gives the product of TDP-4,6-dideoxy-3-keto-glucose (58).[53]

The two proposals that have been described above are largely speculative and are only two among several possibilities. The yeast two-hybrid system failed to confirm the formation of a DesI-DesII complex *in vivo*.[135] However, it is still too early to eliminate the possibility that DesI and DesII work together to achieve deoxygenation at C-4. For the sake of simplicity, it is worthwhile to note that whether DesI and DesII act together or independently, the overall reaction from TDP-6-deoxy-4-keto-glucose (**3**) to TDP-4,6-dideoxy-3-keto-glucose (**58**) is a reductive one. Consequently, the amino donor that regenerates PMP must be oxidized. If the amino donor is L-glutamate, the oxidized product should be α-ketoglutarate.

C-O Bond Cleavage/Formation at the C-1 of Hexoses (Figure 1-14)

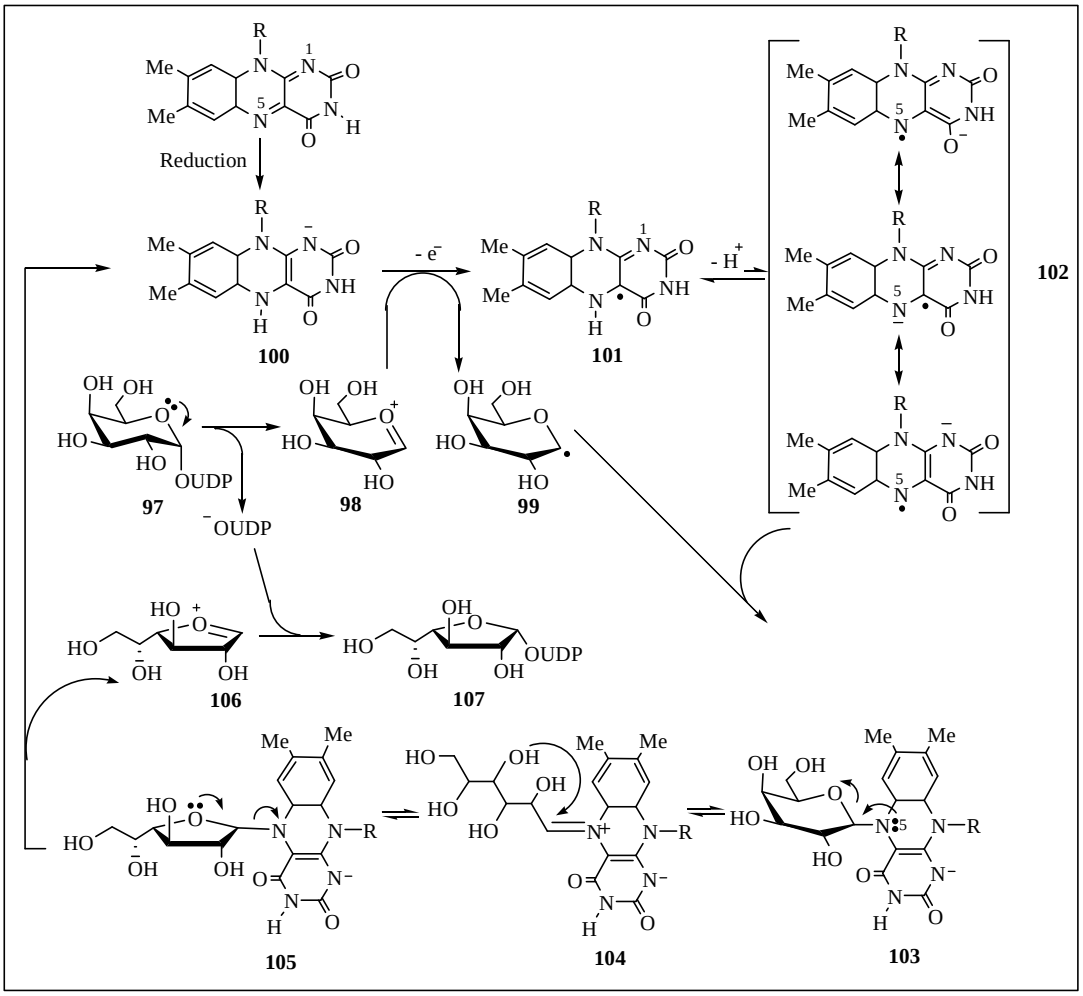


Figure 1- 14. UDP-galactopyranose mutase catalyzed C-O bond cleavage at C-1. Adapted from Huang, Z. et al. Bioorganic Chemistry 2003, 31, 494.

Sugars without an NDP group attached to the anomeric carbon exist in solution in an equilibrium between the linear form and the cyclized form. Thus, bond cleavage between the anomeric carbon and the oxygen from the C5 hydroxyl group of a pyranose or from the C-4 hydroxyl group of a furanose occurs quickly and easily without need of an enzyme. However, once sugars are converted to NDP derivatives, the sugar ring

becomes more stabilized and usually does not reopen during the course of biosynthetic processing. But C-O bond cleavage at the anomeric carbon, usually accompanied by ring contraction, does occur in a few rare cases through the actions of certain enzymes. The conversion from UDP-D-galactopyranose (**97**) to UDP-L-galactofuranose (**107**) catalyzed by the FAD-dependent UDP-galactopyranose (UDP-Galp) mutase is one example.[73,74] Galactofuranose is an abundant component of the cell surface biomolecules of many microorganisms, including bacterial *O*-antigens, fungal exopolysaccharides, protozoal glycoproteins, and the cell walls of mycobacteria.[136] UDP-Galp mutase-catalyzed UDP-galactofuranose formation is mechanistically interesting. When the enzyme is reduced by sodium dithionite, its catalytic efficiency increases significantly. However, since the overall reaction exhibits no net change in the redox state of the substrates involved, it is puzzling how enzyme-bound FAD plays an active role in catalysis. The mechanistic properties of FAD have been studied extensively and are relatively well characterized. FAD can accept and deliver either one electron or two electrons as an enzyme cofactor. As a result, an FAD-catalyzed reaction can go through either a radical or an anion intermediate. To study the catalytic properties of UDP-Galp mutase, the enzyme was reconstituted with deaza-FADs. It was found that the mutase reconstituted with FAD or 1-deaza-FAD had comparable activities, while that reconstituted with 5-deaza-FAD was catalytically inactive.[74] Because 5-deaza-FAD can only catalyze a two-electron process while FAD and 1-deaza-FAD can undergo both concerted two-electron and stepwise one-electron reactions, the loss of the activity of the 5-deaza-FAD reconstituted enzyme supports the following proposed radical mechanism:[74] FAD is first reduced to $FADH^-$

(100) by dithionite or other reducing agent in the actual system. In a likely basic environment, the UDP of the substrate UDP-D-galactopyranose (97) is repelled by the electron donating effect from the ring-oxygen and an oxocarbenium (98) is formed. The two-electron reduced $FADH^-$ (100) then transfers one of the electrons to the oxonium ring to form a radical at C-1 (99), with concomitant conversion of $FADH^-$ to the radical form $FADH^\bullet$ (101). The radical cofactor then loses one proton to form an anion radical delocalized on the isoalloxazine ring (102). The radicals from the substrate and the cofactor are then quenched through formation of a covalent bond between the anomeric carbon and the N-5 of the isoalloxazine ring (103). Next, the C-O bond at C-1 of the sugar ring is broken by the electron donating effect of N-5 to generate an imine intermediate (104). The nucleophilic hydroxyl group at C-4 then forms a C-O bond with the electrophilic carbon of the ketimine to achieve the ring contraction (105). Finally, the UDP is reattached to the anomeric carbon.

The results from the pH dependence studies are consistent with the above proposal. The activity of the UDP-Galp mutase that had been reconstituted with FAD increased considerably at high pHs.[74] The basic environment may facilitate the departure of the UDP group and the deprotonation of the cofactor during the catalysis.

C-O Bond Cleavage at the C-2 of Pentose

The deoxygenation reactions that have been described thus far have one thing in common: the carbon where the C-O bond cleavage occurs has been activated by the formation of a nearby keto group. The activation has either been achieved in formation

of the substrate, or is attained during the course of the catalysis. The 4-keto group activates the α-carbon and β-carbon in C-2, C-3 and C-6 deoxygenation, while to accomplish deoxygenation at C-4, the 4-keto group is converted to a 4-amino by a transaminase to facilitate later bond cleavage. Different from these courses, deoxygenation at the C-2 of pentose is not pre-activated by any functional group formation on the pentose ring. The product, 2-deoxy D-ribose, is used as the building block for DNA synthesis, and is also required for all DNA repairs. Since DNA has existed since the early days of life on earth, it is obvious that the C-2 deoxygenation of pentoses has a tremendously long history. Thus, enzymes catalyzing this reaction, which are called ribonucleotide reductase (RNR), have had considerable time to refine their catalytic mechanisms. Interestingly, the several classes of RNRs known exhibit little sequence similarity, and follow several distinct mechanisms of C-2 deoxygenation of pentose. RNRs that have been identified thus far can be grouped into four classes.[76,77] The classification is based on the cofactor requirements and the mechanism of formation of the thiyl radical, which is required to initiate unactivated C-O bond cleavage (Figure 1-15). Class I RNR enzymes contain a non-heme diiron center, which through complicated interactions with O_2 is able to produce an intermediate tyrosine radical. Homolytic cleavage of the S-H bond of a conserved cysteine residue quenches the tyrosine radical and simultaneously generates the necessary thiyl radical. Class II RNRs use adenosylcobalamin, a vitamin B_{12} derivative, to catalyze the homolytic cleavage of the C-H bond in order to form the thiyl radical. The homolytic cleavage of the Co-C bond in the cofactor generates a 5'-adenosyl radical that abstracts a hydrogen from the active site

cysteine. Class III RNRs contain an iron-sulfur center that is sensitive to O_2, and a SAM molecule that is proposed to bind in the vicinity of the iron-sulfur center. One electron equivalent is transferred through the iron-sulfur center to induce reductive cleavage of the SAM molecule to generate a molecule of methionine and a 5'-adenosyl radical. The formation of the 5'-adenosyl radical induces the production of an enzyme localized

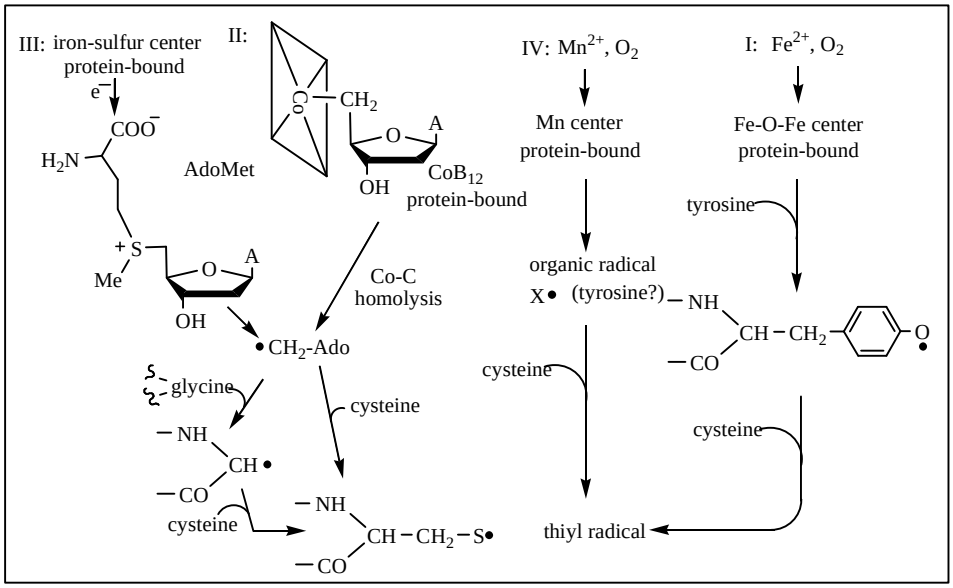


Figure 1-15 . Four classes of RNRs all generate active thiyl radical, but differ from each other by the cofactor requirements. Adapted from Kolberg, M. et al. Biochim Biophys Acta. 2004, 1699(1-2), 1.

glycyl radical that is quenched by the hydrogen radical from S-H homolytic cleavage of the active site cysteine. As a result, the thiyl radial, which is necessary to commence the following catalysis, is generated. Class IV RNRs, which contain a protein bound Mn^{2+}

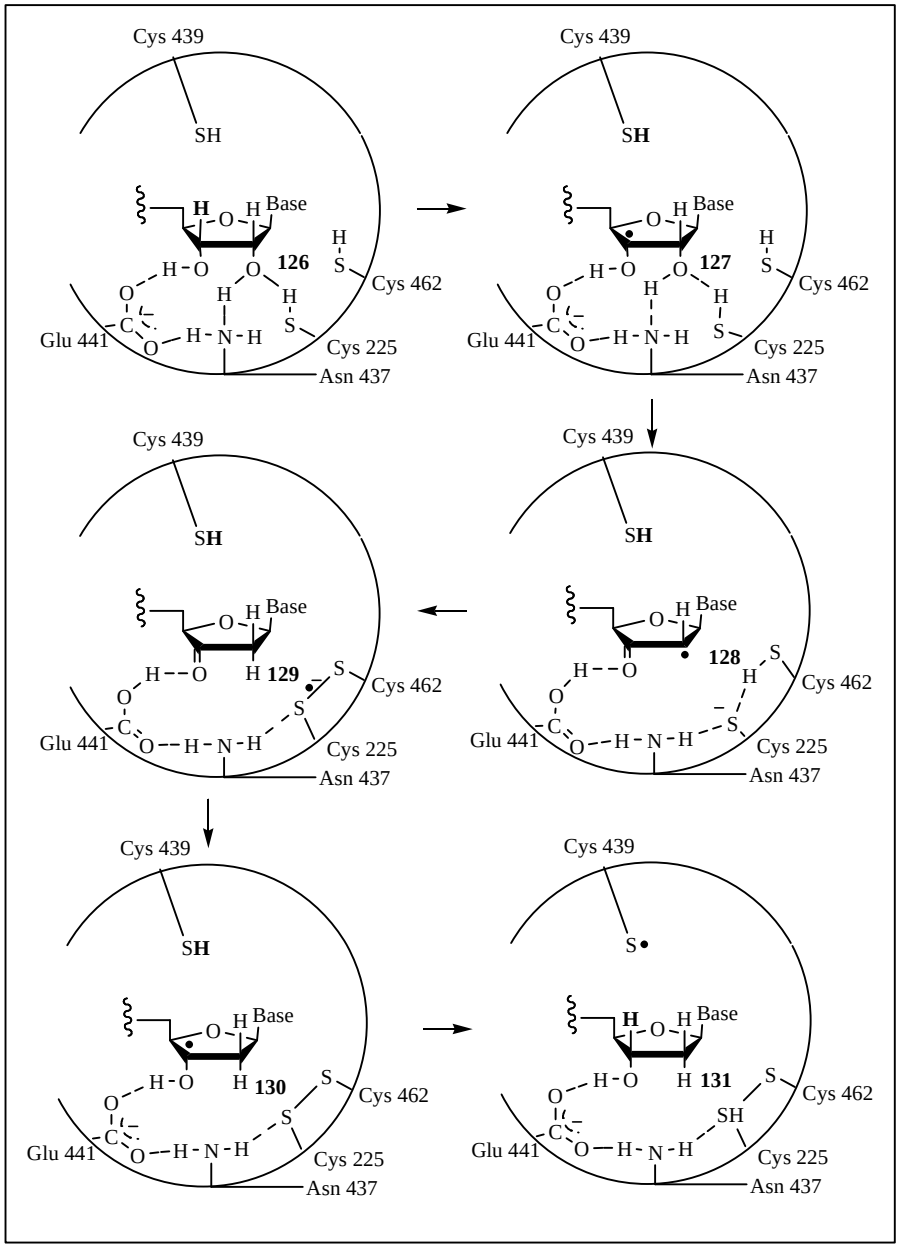


Figure 1-16 . RNRs catalzyed C-2-deoxygenation of pentose. Adapted from Follmann, H. Chem Soc Rev. 2004, 33(4), 225.

center, are less characterized. It is suggested that due to a complicated interaction with O_2, an intermediate radical, possibly a tyrosine radical, is generated. This radical then

abstracts the hydrogen from the S-H bond of an active site cysteine to yield the necessary thiyl radical.

Once the thiyl radical is formed, the enzymes from the four different classes display a similar chemical strategy (Figure 1-16).[77,78] The thiyl radical first abstracts the hydrogen from the C-3 of the pentose to form a radical on the sugar ring. Next, 3-keto formation assisted by the active site glutamate through hydrogen bonding drives the elimination of the hydroxyl group at the C-2, which leaves as a H_2O molecule. At the same time, the single electron at C-3 is relocated to the C-2. Next, the radical at the C-2 abstracts a hydrogen from one of the two other active site cysteines. It should be noted that neither of these cysteines is the one that generates the thiyl radical and abstracts the hydrogen from the C-3. Disulfide bond formation between these two cysteines generates a protein anion radical. The extra electron, along with a proton, reduces the keto group at C-3 to a hydroxyl group. Consequently, a C-3 radical is again generated. The C-3 radical then abstracts the hydrogen that has been taken by the primary active site thiyl radical to complete the C-2 deoxygenation and regenerate the thiyl radical. However, the protein disulfide bond formed during catalysis must be reduced in order to start another catalytic cycle.[77]

4. C-N BOND FORMATION IN DEOXYSUGAR BIOSYNTHESIS

Aminosugars, a class of deoxysugars whose hydroxyl group is replaced with an amino group or an amino derivative, have been found to be constituents of glycoproteins, glycolipids and a variety of secondary metabolites. They have been shown to play a pivotal role in determining the efficacy and specificity of numerous clinically relevant

natural products.[75] Methylation and acetylation are common enzymatic modifications of the amino group. The C-N bond thus generated is not on the backbone of the sugar ring. The derivations usually occur by nucleophilic addition of the amino group to an electrophilic carbon that is either the methyl cation of SAM or the carbonyl center of an acetyl group. In contrast, C-N bond formation involving a backbone carbon usually occurs by a transamination reaction. That is, the keto group is replaced with an amino group by the catalysis of a coenzyme B_6-dependent enzyme.

C-N Bond Formation by Transamination at C-3 and C-4 of Hexose (Figure 1-17)

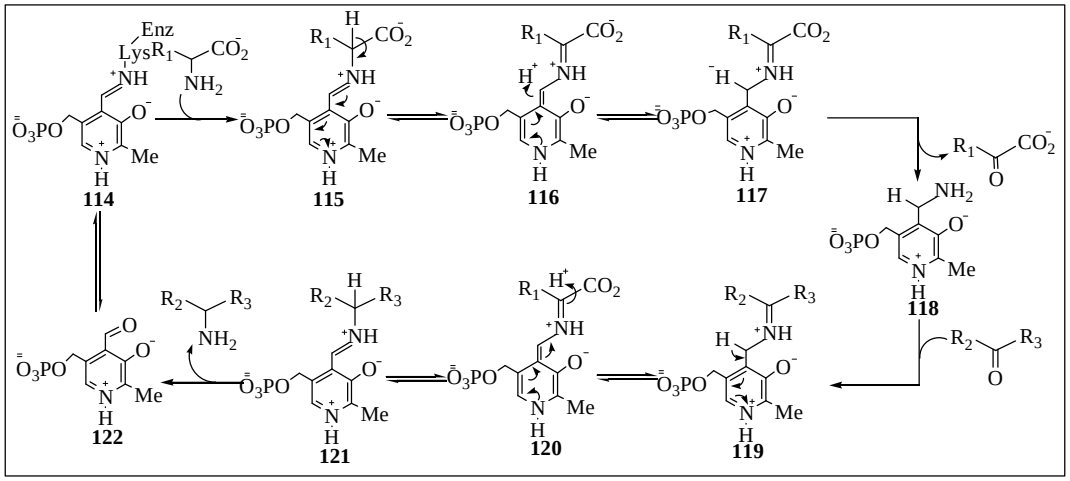


Figure 1-17 . Coenzyme B_6 catalyzed transamination.

Although very few transamination reactions in the aminosugar biosynthesis have been characterized, it is probable that these reactions follow a mechanism similar to other aminotransferase catalyzed reactions. The overall catalysis can be divided into two discrete half-reactions. In the first half-reaction, an amino donor enters the active site of

the enzyme to form a Schiff base with the coenzyme PLP that has been covalently linked to a conserved lysine residue (115). An α-proton is then abstracted and the electron equivalent from the generated carbanion is relocated and deposited in the pyridine ring of the ketimine intermediate (116). After electron rearrangement, the C-4' carbon of the intermediate is protonated (117). Hydrolysis of the ketimine yields an α-keto acid and the PLP coenzyme is converted to the PMP form (118). In this half reaction, the coenzyme undergoes a two-electron reduction from an aldehyde to form the primary amine in PMP. The second half-reaction begins with Schiff base formation between the $2e^-$ reduced PMP and the keto sugar substrate (119). The hydrogen at C-4' of the ketimine intermediate is then abstracted by a base. The electron on the generated C-4' anion is relocated in the pyridine ring of PMP (120). Next, tautomerization of the PMP-hexose intermediate protonates the α-carbon of the sugar and regenerates the pyridinium ring of PMP (121). Lastly, hydrolysis releases the aminosugar as the final product and the coenzyme is converted back to its PLP form (122). As detailed analysis reveals, the second half reaction is simply the reverse of the first half reaction. The amino group is provided by the amino donor, carried by the coenzyme and then transferred to the sugar substrate. In contrast, the amino donor is oxidized to the keto product.

Some enzymes deduced from gene sequences in biosynthetic clusters of aminosugars contain coenzyme B_6 binding motifs and show significant sequence similarity to known aminotransferases.[137] The functions of some of them have been confirmed by either *in vivo* gene deletion studies or *in vitro* biochemical characterization. C-3 aminotransferases have been shown to be regiospecific, being unable to take 4-keto

sugars as substrates.[49,72] It is speculated that C-4 aminotransferases are also unable to take 3-keto sugars as substrates. It is not surprising to observe that the sequences of C-3 aminotransferases and C-4 aminotransferases are only distantly related.

C-N Bond Formation at C-2 of Hexose (Figure 1-18)

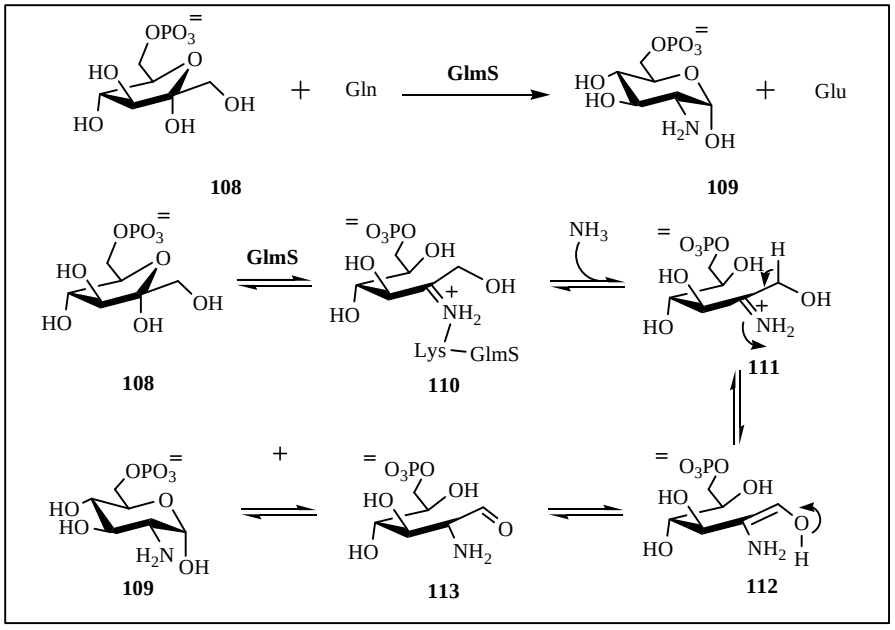


Figure 1- 18. GlmS catalyzed C-N bond formation. Adapted from Badet-Denisot, M. A. et al. Arch. Biochem. Biophys. 1997, 337, 129.

The well-studied glucosamine-6-phosphate synthase catalyzes the conversion from fructose 6-phosphate to glucosamine 6-phosphate using glutamine as the nitrogen donor.[138] Glucosamine 6-phosphate is a precursor of uridine diphospho-*N*-acetylglucosamine from which other C-2 aminosugars are generally derived. In addition to C-N bond formation at C-2, the overall reaction includes a ring expansion from a furanose to a pyranose. Glucosamine 6-phosphate synthase (GlmS) is an enzyme that

catalyzes two distinct enzymatic reactions. The glutaminase domain catalyzes the hydrolysis of glutamine to yield glutamate and enzyme-sequestered ammonia. The newly formed NH_3 is then transferred to the fructose 6-phosphate acceptor domain. It is believed that the opening of the furanose ring of the substrate begins with the formation of a Schiff base between the substrate sugar and an active site lysine of GlmS (110). The NH_3 transferred from the glutaminase domain then reacts with the enzyme-substrate ketimine intermediate to form a substrate-substrate Schiff base (111). Subsequent isomerization gives rise to an aldose product (109).[139-141]

The stereochemical course of the ketose/aldose isomerization in the glucosamine biosynthesis has been studied. By using both (1*R*)-[1-^{3}H] and (1*S*)-[1-^{3}H]fructose-6-phophate as substrates, GlmS was shown to catalyze the stereospecific exchange of the *pro-R* C-1 proton of the sugar. The vast majority of (1*R*)-^{3}H was released into medium.[142]

C-N Bond Formation on the Amino Group(Figure 1-19)

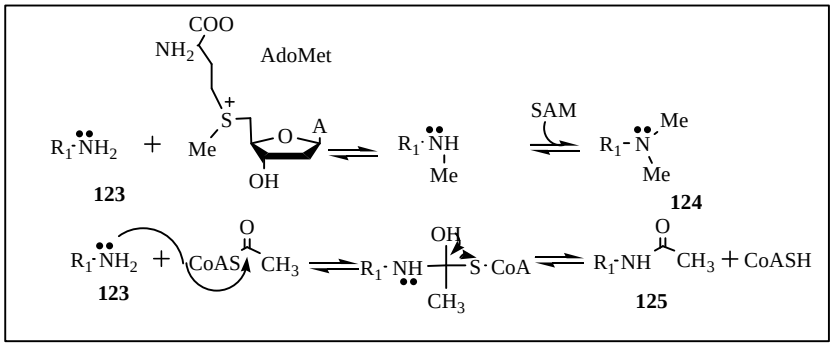


Figure 1-19 . Amino derivations catalyzed by acyltransferase and methyltransferase.

The amino groups of many aminosugars are further modified to fine-tune their biological activities. *N*-methylation and *N*-acetylation are the two common types of amino group derivations that have been demonstrated in many systems with purified enzymes. It is well accepted that the modifications are facilitated by the nucleophilic nature of the amino groups of the sugar, although the detailed mechanism of this type of reactions has not been demonstrated. *N*-methylation can go through an S_N1-like reaction pathway where SAM, the typical methyl donor, breaks down to form a molecule of methionine and a methyl cation, with subsequent nucleophilic addition to make the linkage between the methyl group and the amino group. Alternatively, the nucleophilic addition of the amino group and the leaving of the methionine can happen concomitantly in an S_N2-like reaction. In *N*-acetylation reaction, it is believed that the derivitization is completed by nucleophilic addition of the amino group to the carbonyl center of acetyl-CoA. A tetrahedral intermediate is formed first, followed by elimination to release the free CoA and the acetylated amino group. The reaction course of the *N*-acetylation may be analogous to the nucleophilic replacement reactions known to occur with the carbonyl group.

5. Summary

The C-O bond cleavage reaction is a fundamental process in synthesizing deoxysugars. In recent years, progress has been made in understanding the enzymatic mechanisms of deoxygenations at the various carbons on the sugar unit. C-2

deoxygenation occurs through a β-elimination/reduction route. Deoxygenation at C-3 may undergo α-elimination or step wise reduction, with PMP playing an important role in either case. C-4 deoxygenation may involve a [4Fe-4S] center reduced SAM radical. C-1 deoxygenation accompanied by ring contraction may involve an FAD-dependent radical mechanism.

Compared with the C-O bond cleavage, the C-N bond formation is relatively straightforward, occurring either by nucleophilic addition or transamination. The seemingly complicated chemistry of glucosamine-6-phosphate synthase is, in essence, still a transamination reaction at C-2 of the sugar.

Deoxysugars have probably existed as components of the cell surfaces of various organisms throughout evolution. Due to this long evolutionary history, it is reasonable that the mechanisms involved in the C-O bond cleavage are diverse. Thus, it is logical to keep an open mind for the discovery of other possible mechanisms of C-O bond cleavage. A comprehensive understanding of many of the mechanisms proposed thus far must await more detailed studies as well.

6. **Thesis Statement**

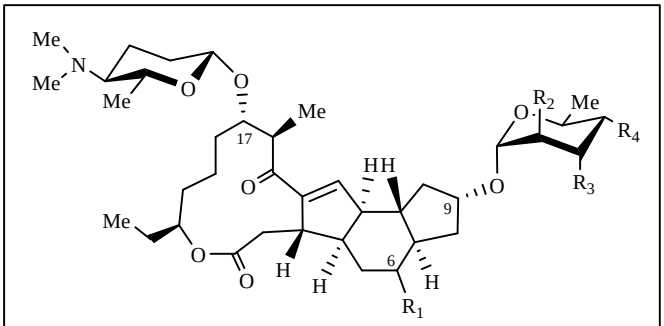
In spite of the great progress that has been made in the last twenty years, the explicit characterization of deoxysugar biosynthetic pathways remains limited. Many of the proposed pathways remain speculative. In this thesis, detailed studies of the biosynthesis of TDP-D-forosamine are carried out. Overexpression, purification, and characterization of each enzyme in the biosynthetic pathway of forosamine have been achieved. The enzymatic functions and cofactor requirements of each enzyme have been

unambiguously established by product separation and characterization. Comprehensive biochemical studies of each enzyme have been performed. The enzymatic reactions include C-O bond cleavage at C-2, C-3, C-4 and C-6. New evidence has been obtained to support the C-2 β-elimination/reduction mechanism. An aminotransferase catalyzing the C-4 transamination reaction has been identified. This work also gives a firm explanation why the forosamine biosynthetic gene cluster lacks an E_3 reductase homologue involved in C-3 deoxygenation, as well as demonstrates a reaction system for C-3 deoxygenation that is different from the well-established E_1/E_3 pair. In addition, the methyltransferase identified in this study is the first 4-amino-*N*,*N*-dimethyltransferase that has ever been characterized. The step-wise methylation catalyzed by this methyltransferase is established as well. Together, studies carried out in this thesis have significantly enriched our knowledge of deoxysugar biosynthesis. Forosamine is the first 2,3,4,6-tetradeoxysugar whose biosynthetic pathway has been completely characterized.

Chapter 2. Characterization of SpnO, a TDP-6-deoxy-4-keto-D-glucose 2,3-dehydratase and SpnN, a TDP-2,6-dideoxy-3,4-diketo-D-glucose 3-ketoreductase

1. INTRODUCTION

The spinosyns are novel secondary metabolites produced by *Saccharopolyspora spinosa*. Spinosyns are comprised of a tetracyclic macrolide and appendages D-forosamine and L-rhamnose, differing from each other by the degree of methylation on the polyketide or the L-rhamnose moiety (Figure 2-1). The two major spinosyn



Factor	R_1	R_2	R_3	R_4
A	H	OCH_3	OCH_3	OCH_3
D	CH_3	OCH_3	OCH_3	OCH_3
H	H	OH	OCH_3	OCH_3
J	H	OCH_3	OH	OCH_3
K	H	OCH_3	OCH_3	OH
P	H	OCH_3	OH	OH

Figure 2-1 . Structures of spinosyns. Forosamine and tri-O-methylrhamnose are attached at positions 17 and 9, respectively.

components in *Saccharopolyspora spinosa* extracts are spinosyn A and spinosyn D, differing from each other by the presence or absence of a single methyl substituant at position 6 of the polyketide. Spinosad, a combination of spinosyn A and spinosyn D, is marketed by Dow AgroSciences for control of agricultural insect pests. Spinosad is highly effective against target insects and has an excellent environmental and mammalian toxicological profile.[143-145] Incorporation studies with ^{13}C-labeled acetate, propionate, and methionine established that spinosyns are assembled by a polyketide pathway, and that the two *N*-methyl groups of forosamine and the three *O*-methyl groups of tri-*O*-methylrhamnose are derived from SAM.[146]

The spinosyn biosynthetic gene cluster has been identified and sequenced (Figure 2-2).[147] The macrolide of spinosyn differs from common type-I polyketide macrolide in

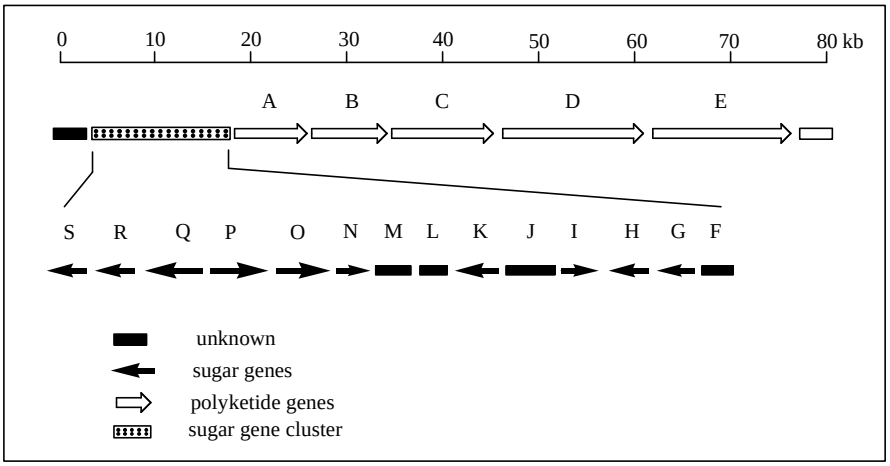
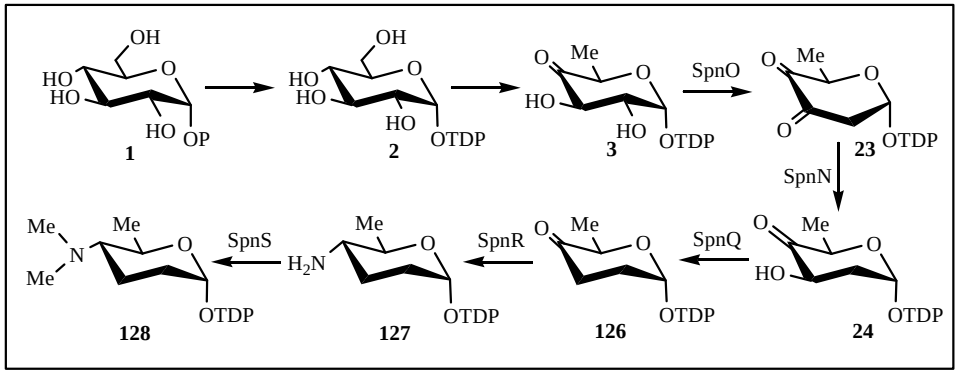


Figure 2- 2. Organizations of the spinosyn biosynthetic gene cluster.

that it is tetracyclic instead of being monocyclic. Identification of the genes involved in these C-C bond formation steps helped in locating the whole biosynthetic gene cluster.[147] To aid the cloning and analysis, a gene transfer system that employed conjugation from *E. coli* was developed.[148] Using conjugation, cosmids containing large inserts of *Saccharopolyspora spinosa* DNA were introduced into blocked mutants of *S. spinosa*. These inserts were efficiently inserted into the chromosome by homologous recombination, and the effects of the introduction studied.[148] By gene disruption and bioconversion analysis, the boundaries of the spinosad biosynthetic gene cluster were established. Interestingly, the TDP-L-rhamnose biosynthetic genes were not found in the sequenced gene cluster. This is probably because TDP-L-rhamnose is also an important primary metabolite and thus the organism has multiple routes to synthesize this sugar. From sequence analysis, the genes *spnH*, *spnI*, and *spnK* are proposed to encode rhamnose *O*-methyl transferase while *spnG* is proposed to encode rhamnosyltransferase. In addition, the enzymes encoded by genes *spnO*, *spnN spnQ*, *spnR* and *spnS*, are proposed to participate in TDP-D-forosamine biosynthesis. The five genes are adjacent to each other and located close to the PKS region. However, the genes encoding the thymidylyltransferase and TDP-glucose-4,6-dehydratase were not found in the vicinity. The biosynthesis of TDP-D-forosamine might recruit these two enzymes directly from the cellular pool since these two enzymes are widely used in many deoxysugar biosynthetic pathways.

The biosynthetic pathway of forosamine has been proposed based on sequence analysis (Scheme 2-1). It is proposed that the biosynthesis of TDP-D-forosamine begins

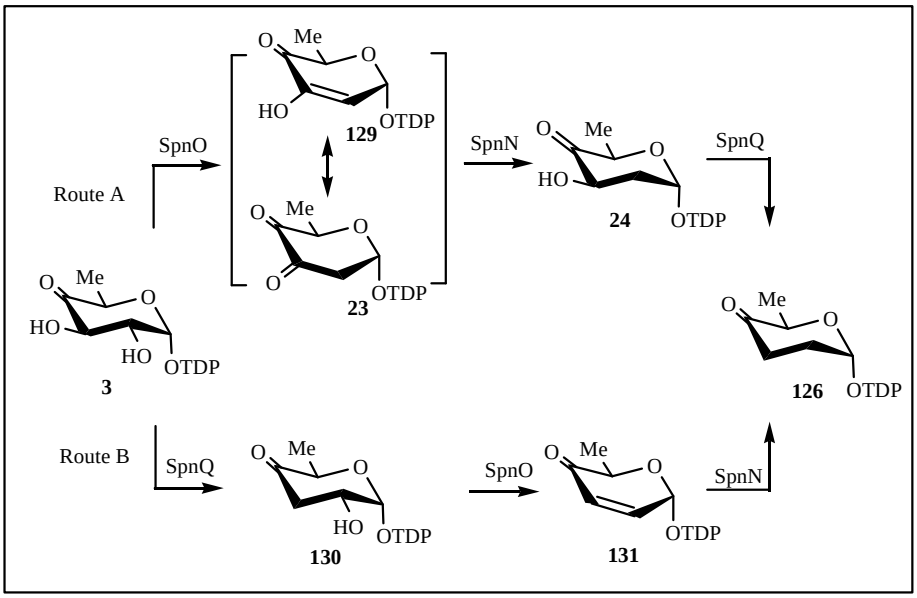
with glucose-1-phosphate (**1**). TDP-6-deoxy-4-keto-glucose (**3**) is synthesized by a glucose-1-phosphate thymidylyltransferase and a TDP-glucose-4,6-dehydratase from the cellular pool. Next, the 2-hydroxyl group of TDP-6-deoxy-4-keto-glucose is proposed to be removed by action of the enzymes encoded by *spnO* and *spnN*. The TDP-2,6-dideoxy-4-keto-glucose (**24**) is then deoxygenated at the C-3 by the *spnQ*-encoded enzyme to give **126**. The two enzymes encoded by *spnR* and *spnS* are proposed to catalyze transamination of C-4 to give **127**, and *N*,*N*-dimethylation of the resulting 4-amino sugar (**127**), respectively. The glycosyltransferase that attaches forosamine (**128**)



Scheme 2-1. Proposed TDP-forosamine biosynthetic pathway. The first two steps in the pathway are suggestedto be catalyzed by enzymes from the cellular pool, while the other steps are catalyzed by enzymes encoded in the spinosyn biosynthetic gene cluster.

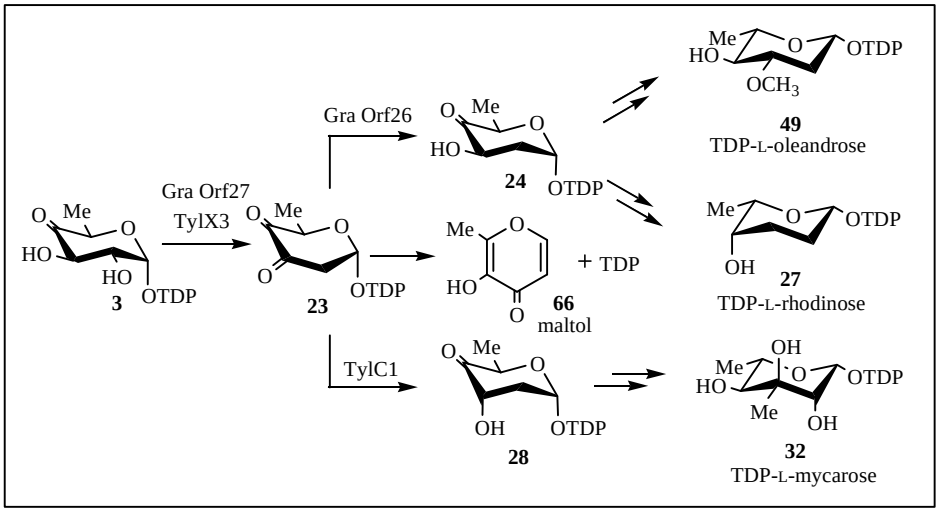
to the polyketide is probably encoded by *spnP*. As has been explained in Chapter 1, few deoxysugar biosynthetic pathways have been completely characterized in spite of the progress made in assigning enzymatic functions to the gene products of deoxysugar biosynthetic gene clusters. Most of the biosynthetic pathways in the literature are still speculative. The correlation between each enzyme and the specific reaction it catalyzes

in the pathway is rarely unambiguously established. Forosamine is the most highly deoxygenated sugar whose biosynthetic pathway has been studied to date. The hydroxyl groups at C-2, C-3, C-4, and C-6 in the hexose have been replaced with hydrogen, hydrogen, *N*,*N*-dimethyl amino group, and hydrogen, respectively. However, the assigned reaction sequence is only tentative. It is reasonable to consider that the SpnS catalyzed reaction would occur after the SpnR catalyzed reaction since SpnS is proposed to be an *N*,*N*-dimethyltransferase and the substrate must have installed an amino group. Moreover, the SpnR reaction should occur later than the C-2 and C-3 deoxygenations because the deoxygenation reactions at the C-2 and C-3 require the presence of the C-4



Scheme 2-2. Two possible pathways to obtain TDP-2,3,6-trideoxy-4-keto-glucose. In Route A, the C-2 deoxygenation occurs before the C-3 deoxygenation. In Route B, vice versa.

keto functional group to activate the substrate. However, the order of the C-2 and C-3 deoxygenations is uncertain as none of the biosynthetic pathways of 2,3-dideoxysugars have been previously characterized (Scheme 2-2). Thus, TDP-6-deoxy-4-keto-glucose (**3**) might first undergo C-3 deoxygenation (**130**) followed by C-2 deoxygenation. In addition, the participation of SpnO and SpnN in C-2 deoxygenation is based solely on their homology to TylX3/TylC1 and Gra Orf27/Gra Orf26 whose functions have been demonstrated *in vitro*,[70,71] and thus must be confirmed. The respective aminotransferase and methyltransferase activities of SpnR and SpnS are also yet to be demonstrated. Moreover, the C-3 deoxygenation proposed to be catalyzed by SpnQ is intriguing. In this chapter, the study of the first two enzymatic steps in TDP-D-forosamine biosynthesis, proposed to be catalyzed by SpnO and SpnN, will be discussed.



Scheme 2-3. Biosynthesis of TDP-L-oleandrose and TDP-L-rhodinose requires Gra Orf27/Gra Orf26, while biosynthesis of TDP-L-mycarose requires TylX3/TylC1.

BLAST search of the *spnO* gene product through the National Center for Biotechnology Information (NCBI) data base has found more than 30 homologs of SpnO with identity higher than 25%. The functions of two of them have been characterized by *in vitro* studies (Scheme 2-3). TylX3 (43.8% identity to SpnO) was demonstrated to catalyze 2,3-dehydration in the biosynthesis of TDP-L-mycarose (**32**), one of the deoxysugar components in antibiotic tylosin that is produced by *Streptomyces* fradiae.[71] Gra Orf27 (41.8% identity to SpnO) catalyzes the same reaction in the biosynthesis of TDP-L-oleandrose (**49**) and TDP-L-rhodinose (**27**), deoxysugar components of the antibiotics granaticin and granaticin B that are produced from *Streptomyces violaceoruber* Tü22[70]. BLAST search of the *spnN* gene product gives numerous proteins with significant similarity to SpnN. Most of them are ketoreductases, while others are assigned as epimerases. The sequence similarity between ketoreductase and epimerase can be rationalized by the common use of NAD(P)(H) cofactor. Among these homologs, the enzyme Gra Orf26, which has high sequence identity to SpnN (47.7%), has been shown to be a ketoreductase in the biosynthetic pathway of TDP-L-oleandrose (**49**) or TDP-L-rhodinose (**27**) (Scheme 2-3).[70] Interestingly, TylC1 is only distantly related to SpnN. It should be noted that Gra Orf27 and Gra Orf26 are found in the same gene cluster. Gra Orf26 was demonstrated to convert the product of Gra Orf27 (**23**) to TDP-2,6-dideoxy-4-keto-glucose (**24**) with the reintroduced C-3 hydroxyl group in an equatorial position. In summary, based on the sequence similarity to characterized enzymes, SpnO and SpnN are proposed to catalyze deoxysugar 2,3-dehydration and 3-

ketoreduction, respectively. It is interesting to note that the two genes *spnO* and *spnN* are physically next to each other in the *S. spinosa* genome.

2. EXPERIMENTAL PROCEDURES

Materials. *Saccharopolyspora spinosa* genomic DNA, used as a PCR template for amplification of *spnO* and *spnN*, was previously prepared by Dr. Kent Zhao. *E. coli* DH5α cells were purchased from Bethesda Research Laboratories (Muskegon, Michigan). Vector pET17b(+), pET24b(+), pET28b(+), and pET41a(+) were purchased from Novagen (Madison, WI). Overexpression host cell *E. coli* BL21(DE3), BL21(DE3) pLysS, HMS174(DE3), Rosetta^{TM}2(DE3), and RosettaBlue^{TM}2(DE3) were also obtained from Novagen. Overexpression host BL21-CodonPlus(DE3)-RP was purchased from Stratagene (La Jolla, CA). Components of culture media were purchased from Difco (Detroit, MI). *pfu* DNA polymerase was obtained from Stratagene. Qiaquick Linear DNA Purification Kit and Qiaprep DNA Miniprep Kit were purchased from Qiagen (Valencia, CA). The 1kb DNA ladder and agarose were from Gibco BRL (Grand Island, NY). Oligonucleotides used in the PCR amplification of desired inserts were prepared by Invitrogen (Carlsbad, CA) and used without further purification. Restriction endonucleases were from either Invitrogen, Gibco BRL, or Promega (Madison, WI). Ni-NTA resin for purification of His-tagged proteins was a product of Qiagen. Bio-Gel P2 extra-fine resin was purchased from BioRad (Richmond, CA). Small amount of solution was disinfected by filtering through Sterile Syringe Filter purchased from Corning Inc (Coring, NY). Protein markers were obtained from Sigma (St. Louis, MO). Mini trans-

blot assembly apparatus for Western blot and non-fat dry milk were purchased from BioRad. Also used for Western blot were nitrocellulose Hybond-C Extra membrane from Amersham (Piscataway, NJ), and each from Sigma, monoclonal anti-polyhistidine clone His-1, anti-mouse IgG, and alkaline phosphatase conjugate. For Western blot detection, nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) were both obtained from Promega. NADH, NADPH, TMP, morpholine, DCC, and 1*H*-tetrazole were purchased from Sigma-Aldrich. All other chemicals were analytical grade or the highest quality commercially available. Sequence analysis and comparison were performed using Vector NTI software and BLAST network service with the NCBI computing system.

General Equipment. Ultraviolet-visible data were recorded with a Beckman DU-650 spectrophotometer. The pH values were obtained using a Corning pH meter 240. The cell membrane disruption was done using Fisher 550 Dismembrator from Fisher Scientific (Pittsburg, PA). The DNA-100 submarine gel equipment from Pharmacia Biotech (Uppsala, Sweden) powered by either a FB600 or FB 135 supply from Fisher Scientific were used for all agarose gel electrophoresis. DNA gel documentation was performed with camera FBTIV-88 and ID (version 3.5) software from Kodak (Rochester, New York). DNA gels were illuminated by Transilluminator from Fisher Scientific. Polymerase chain reactions were carried out in either Perkin Elmer Cetus DNA Thermal Cycler (Norwalk, CT) or Mastercycler Gradient Thermal Cycler from Brinkmann Instruments (Westbury, NY). The Mini-Protein II electrophoresis system for SDS-PAGE analysis was manufactured by Bio-Rad. Polyacrylamide gels were dried using a Bio-Rad

Gel Air Dryer. HPLC separation was achieved on a HP1090 Liquid Chromatograph from Hewlett Packard (Rolling Meadow, IL). Centrifugation was performed with an Avanti J-25 or J-10 machine from Beckman (Arlington Heights, IL), while microcentrifugations were done with an Eppendorf 5415D from Brinkmann Instrument.

Preparation of Competent Cells. Cells were made competent according to a slightly modified $CaCl_2$ method.[154,155] A single fresh colony of an appropriate *E. coli* strain was used to inoculate 5 mL of LB medium and the resulting culture was grown overnight at 37 °C with shaking at 250 rpm. A 400 μL aliquot of this overnight culture was used to inoculate 100 mL of LB medium in a sterile Erlenmeyer flask. When the cell growth at 37 °C reached an OD_{600} of about 0.4, the culture was apportioned into two prechilled polypropylene tubes (Falcon 2070) and left on ice for 10 min. After centrifugation at 1,600 g for 7 min, the cell pellets were gently resuspended in 5 mL of ice-cold $CaCl_2$ solution (60 mM $CaCl_2$, 15% glycerol, 10 mM PIPES, pH 7.0, disinfected by filtering through sterile syringe filter). Another centrifugation step at 1,100 g, 4 °C for 7 min was performed and the cells were then resuspended in 5 mL of ice-cold $CaCl_2$ solution. After incubation on ice for 30 min, the cells were spun down for 5 min at 1,100 g, 4 °C. Finally, the cells were resuspended in 500 μL of the cold $CaCl_2$ solution and aliquotted in 20 μL portions into prechilled microcentrifuge tubes and frozen at –80 °C or used directly for transformation.

PCR Amplification of the spnO Gene. Three oligonucleotide primers were obtained to amplify the *spnO* gene: primer (O-1) 5'-GGGGCCT**CATATG** AGCA-GTTCTGTCGAAGC-3' containing an *Nde*I restriction site (in bold), primer (O-2) 5'-

GAG**AAGCTT**TTATCGCCCCAACGCCCACAAG-3' containing a *Hind*III restriction site (in bold), and primer (O-3) 5'-TAT**AAGCTT**TCGCCCCAACGCCCACAAGC-3' containing a *Hind*III restriction site (in bold). The start primer O-1 is designed to include the first six codons of the *spnO* gene while the halt primers O-2 and O-3 include the last six codons of the *spnO* gene upstream of the *Hind*III restriction site. Two polymerase-mediated amplifications were carried out, both in 0.5 mL microcentrifuge tubes. In the first reaction, to the reaction mixture was added 37.5 μL of double-deionized H_2O, 5.0 μL of *pfu* polymerase buffer (10 x), 4.0 μL of deoxyribonucleotidyltriphosphate mix (2.5 mM each), 1.0 μL of start primer O-1 (10 μM), 1.0 μL of halt primer O-2 (10 μM), 0.5 μL of genomic DNA (about 0.1 μg) from *Saccharopolyspora spinosa* as the amplification template, and 1.0 μL of cloned *pfu* polymerase (2.5 units). The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following thermal cycles: (1) 1 denaturation cycle of incubation at 95 °C for 3 min; (2) 30 amplification cycles of incubation at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 1.5 min; (3) 1 extension cycle of incubation at 72 °C for 10 min. The tube was held at 4 °C prior to being removed from the thermal cycler. The amplification product is the natural *spnO* gene with restriction sites of endonucleases *Nde*I and *Hind*III at the 5'- and 3'- ends, respectively. In the second polymerase-mediated amplification reaction all the other components are the same as in the first set of reaction, except that the halt primer is changed to O-3 of the same concentration. The thermal cycles and the afterward treatment are also the same. The amplification product from the second reaction is the *spnO* gene without the natural

stop codon and with restriction sites of endonucleases *Nde*I and *Hind*III at the 5' and 3' ends, respectively.

Cloning of the spnO Gene. After PCR amplification, standard recombinant DNA techniques were used to clone the genes into appropriate expression vectors. The PCR product mixtures from the two PCR reactions were individually separated from the mineral oil and electrophoresed on a 0.7% agarose gel. The DNA fragments of the correct size were purified with QIAquick Spin Kit. The cleaned products of the two PCR reactions were each double digested with *Nde*I and *Hind*III in appropriate buffer (10 x restriction buffer 2, Invitrogen) at 37 °C overnight. The mixtures from the two digestion reactions, containing the two digested *spnO* genes that are named O-w-stop-diges and O-wo-stop-diges, were loaded on a 0.7% agarose gel for electrophoresis. The desired DNA fragments with the size of the *spnO* gene were excised and purified with QIAquick Spin Kit. Digestion of vectors pET24b(+) and pET28b(+) with *Nde*I and *Hind*III was carried out in an appropriate buffer (10 x restriction buffer 2, Invitrogen) for two hours at 37 °C. The digestion product of O-w-stop-diges was ligated into each of digested vectors pET28b(+) and pET24b(+) using high concentration T4 DNA ligase in 5 x ligase buffer. The reactions were carried out at room temperature for 2 h. Ligation into pET28b(+) created a 60bp in-frame extension of the gene, containing six successive His codons, attached to the 5' end of the gene. When expressed, this clone would produce an *N*-terminal His_6-tagged fusion protein. Ligation of the same insert into pET24b(+) maintains the natural gene which at the time of protein expression would produce the native protein. The digestion product of O-wo-stop-diges was ligated into the digested

vector pET24b(+) using high concentration T4 DNA ligase in 5 x ligase buffer. The ligation product contained a 39bp in-frame extension, containing six successive His codons, attached to the 3' end of the gene, producing a *C*-terminal His_6-tagged fusion protein when expressed.

Transformation of 20 μL of *E. coli* DH5α host cells was carried out with 20 ng of the ligation product. The resulting cells were incubated on ice for 30 min followed by heat shock at 42 °C for 90 s. The cells were then incubated on ice for another 2 min, and mixed with 800 μL of LB medium. The cells were incubated at 37 °C for 45 min with shaking at 200 rpm and then briefly centrifuged (1 min at 13,000 rpm). A total of 700 μL of the supernatant was removed. The small cell pellet was resuspended in the remaining medium and then plated out on LB plates supplemented with 30 μg/mL of the antibiotic kanamycin. The plates were incubated for 20 h at 37 °C before six clonies were picked to individually inoculate small liquid cultures. Plasmid DNA was isolated from 5 mL of each of the cultures that had grown for 12 h at 37 °C by using QIAprep Spin Miniprep Kit. An amount of 40 ng plasmid DNA from each miniprep was digested with restriction enzymes *Nde*I and *Hind*III. Positive clones were identified by the presence of the DNA fragment having the size of *spnO* on the agarose gel after electrophoresis and were further confirmed by DNA sequencing. The plasmid DNA of the positive clones, named pLHON having extra His codons at the 5' end of the natural *spnO* gene, pLHOO maintaining the natural *spnO* gene, and pLHOC having extra His codons at the 3' end of the *spnO* gene, were used to transform *E. coli* BL21(DE3).

PCR Amplification of the spnN-short Gene. This work was completed by Dr. Zhao in our laboratory. Many PCR reactions using various primer pairs were carried out in generating clones for the many trials of expression of the *spnN*-short gene, which was named after the gene deposited in the NCBI data base. Two of them were described here. The start primer (pKZ-8), 5'-GGC**CATATG**CGAAAGCCGGTGCGCATC-3', contains an *Nde*I restriction site (in bold). The ATG sequence at the *Nde*I site could also serve as the start codon for the *spnN* gene. The halt primer (pKZ-12), 5'-GCC**CTCGAG**CTATGTGGACCCGCA-3' incorporated an *Xho*I restriction site (in bold) immediately downstream from the coding sequence. Another halt primer (pKZ-9), 5'-GCC**CTCGAG**TGTGGACCCGCACCGACG-3', also contains an *Xho*I restriction site. However, the stop codon of *spnN* was removed in pKZ-9 to allow six His codons to be added at the 3'-end of the *spnN* gene after the gene is cloned into pET24b(+). Polymerase mediated amplifications were carried out in 0.5 mL microcentrifuge tubes. To one reaction mixture was added 71.0 μL of double-deionized H_2O, 10.0 μL of *pfu* polymerase buffer (10 x), 1.5 μL of deoxyribonucleotidyltriphosphate mix (10 mM each), 10.0 μL of DMSO, 1.5 μL of 80% glycerol, 2.0 μL of pKZ-8 (20 μM), 2.0 μL of pKZ-12 (20 μM), 0.5 μL of genomic DNA (about 0.1 μg) from *Saccharopolyspora spinosa*, and 1.5 μL of cloned *pfu* polymerase (2.5 U/ μL). The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following thermal cycles: (1) 1 cycle of incubation at 95 °C for 1 min; (2) 30 cycles of incubation at 95 °C for 1 min, 55 °C for 1 min, and 70 °C for 3.5 min; (3) 1 cycle of incubation at 72 °C for 10 min. The tubes

were held at 4 °C prior to further manipulations. In the other reaction, the pKZ-12 was replaced with pKZ-9.

PCR Amplification of the spnN-long Gene. In order to amplify the *spnN*-long gene which is in frame with *spnN*-short, but with 12 more nucleotides at the 5'- end of *spn*N-short, another oligonucleotide primer complementary to the sequence at the 5'- end of *spnN*-long was designed. The start primer (pKZ-NL), 5'-GGC**CATATG**ACCAGCTCGATGCGAAAG-3', contains an *Nde*I restriction site (in bold). The polymerase mediated amplification of *spnN*-long had the same components as the amplification reaction of *spnN*-short except that the start primer used was pKZ-NL and the halt primer used was pKZ-9. The thermal cycles and the subsequent treatment were also the same as for *spnN*-short. The resulting PCR product was the *spnN*-long gene without the natural stop codon.

Cloning of the spnN-short and spnN-long Genes. After PCR amplification, standard recombinant DNA techniques were used to clone the genes into appropriate expression vectors. Typically, the PCR product mixtures were separated from the mineral oil and subjected to electrophoresis. The amplified genes were purified with the QIAquick Spin Kit. After digestion with appropriate amounts of *Nde*I and *Xho*I, the mixtures were separately loaded onto a 0.7% agarose gel for electrophoresis. The desired DNA fragments with the correct size of the *spnN* gene were excised and purified with the QIAquick Spin Kit. The DNA inserts were ligated with *Nde*I and *Xho*I-digested vectors using high concentration T4 DNA ligase. Typically, a ligation reaction of 20 μL was carried out with 5.0 units of ligase for 2 h at room temperature.

Transformation of *E. coli* DH5α host cells was carried out with 4 μL of the above ligation mixture. The procedure was the same as that for the *spnO* gene transformation. Positive clones were identified by digesting the plasmid DNA, which had been isolated by QIAprep Spin Miniprep Kit, with restriction enzymes *Nde*I and *Xho*I. The presence of the expected DNA fragments was visualized by agarose gel electrophoresis using 0.7% agarose. Three positive clones will be subsequently discussed: pKZ-N-short-N is *spnN*-short cloned into the vector pET28b(+) with a His_6-tag encoded at the 5' end; pKZ-N-short-C is the *spnN*-short gene cloned into the vector pET24b(+) with a His_6-tag encoded at the 3' end and missing the stop codon of the natural gene; pKZ-N-long-C is *spnN*-long ligated to pET24b(+) with a His_6-tag encoded at the 3' end and missing the stop codon. The plasmid DNA of these three clones was transformed into various types of expression host cells, and expression was attempted in order to find the optimal conditions for protein expression. The expression host cells that were tried include *E. coli* BL21(DE3), BL21(DE3) pLysS, HMS174(DE3), Rosetta^{TM}2(DE3), RosettaBlue^{TM}2(DE3), *E. coli* BL21-CodonPlus(DE3)-RP, as well as others not mentioned here.

Codon Changes of the spnN-long Gene. There is one occurrence of AGGAGG starting from nucleotide 352 in *spnN*-long. Two oligonucleotide primers were designed to mutate the rare Arg codons to Arg codons that are more commonly found in *E. coli*. The upstream primer, 5'-GGGCTGGCC**CGTCGC**AAGAACCTG-3' (pKN-11), contains the mutation from AGGAGG to CGTCGC (in bold) with the rest of the sequence complementary to the *spnN* gene. The downstream primer, 5'-GCAGGTTCTT**GCGACG**GGCCAGC-3' (pKN-12) contains the corresponding changes

(in bold) and is complementary to the start primer. Polymerase mediated mutation was carried out in 0.5 mL microcentrifuge tube. To the reaction mixture was added 39.0 μL of double-deionized H_2O, 5.0 μL of reaction buffer (10 x) of Turbo *pfu* polymerase, 1.0 μL of deoxyribonucleotidyltriphosphate (20.0 mM each), 1.5 μL of start primer pKN-11 (10.0 μM), 1.5 μL of halt primer pKN-12 (10.0 μM), 1.0 μL of template DNA pKZ-N-long-C (50.0 ng), and 1.0 μL of Turbo *pfu* polymerase (2.5 units). The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following thermal cycles: (1) 1 cycle of incubation at 95 °C for 30 s; (2) 16 cycles of incubation at 95 °C for 30 s, 55 °C for 1 min, and 68 °C for 14 min; (3) 1 cycle of incubation at 72 °C for 10 min. The tube was held at 4 °C prior to further treatment.

*Dpn*I restriction enzyme (1.0 μL, 10 U/ μL) was added directly to each amplification reaction below the mineral oil overlay using a small pointed pipet tip. The reaction mixture in the microcentrifuge was spun down for 1 min and immediately incubated at 37 °C for 1 h to digest the parental supercoiled double-stranded-DNA.

The XL1-Blue supercompetent cells, 50 μL, was aliquotted to prechilled microcentrifuge tubes before transformation. The *Dpn*I-treated DNA, 1 μL, in the digestion reaction was added to an aliquot of the supercompetent cells. The resulting cells were incubated on ice for 30 min prior to a 45 s heat pulse at 42 °C. After cooling the reaction on ice for 2 min, 800 μL of LB medium was added to the cooled cells and the resulting culture was incubated at 37 °C for 1 h with shaking at 225 rpm. The cells were spun down and resuspended in 120 μL of LB medium after removal of the extra medium. The culture was then plated out on agar plates containing 30 μg/mL kanamycin. The

transformation plate was incubated at 37 °C for 20 h. Clear clonies with a diameter larger than 0.5 mm were chosen for screening.

Each single colony was used to inoculate 5 mL of LB medium and incubated at 37 °C overnight. An aliquot of 4.5 mL of the culture was spun down and the plasmid DNA was isolated using the QIAprep Spin Miniprep Kit. Several plasmid isolates were subjected to DNA sequencing. The one obtained with the expected mutations was named pKZ-N-long-C-opt, and was used to transform *E. coli* BL21-CodonPlus(DE3)-RP overexpression cells.

Purification of the SpnO Protein. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli-pLHON Cells. An overnight culture of *E. coli* BL21(DE3)-pLHON, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was then lowered to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 15 h at 24 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with one min cooling intervals between bursts. Cellular debris was removed by centrifugation at

35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract, was added 6 mL of Ni-NTA slurry (3 mL of resin) and the resulting mixture was incubated for 2 h at 4 °C with gentle shaking. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle by brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 20 mL of lysis buffer followed by various amounts of wash buffers. The wash buffers and elution buffer used contained the same ingredients as the lysis buffer except that the concentrations of imidazole were altered. The column was washed sequentially with 20 mL of wash buffer containing 20 mM imidazole, 20 mL of wash buffer containing 40 mM imidazole and 15 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 0.5 mL/min. After washing, SpnO protein was eluted with 10 mL of elution buffer containing 250 mM imidazole. Eluted protein was transferred into a dialysis bag and dialyzed against four1.0 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5) for 1 h per liter of buffer. The dialyzed protein was concentrated by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon) and stored at –80 °C.

Purification of the SpnN protein. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli BL21-CodonPlus(DE3)-RP-pKZ-N-long-C-opt Cells. An overnight culture of *E. coli* BL21-CodonPlus(DE3)-RP-pKZ-N-long-C-opt, grown in

LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was then adjusted to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 15 h at 24 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with one min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract was added 8 mL of Ni-NTA slurry (4 mL of resin) and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 30 mL of the lysis buffer followed by various amounts of wash buffers. The wash buffers and elution buffer used contained the same ingredients as the lysis buffer except that the concentrations of imidazole were

altered. The column was washed sequentially with 30 mL of wash buffer containing 20 mM imidazole, 30 mL of wash buffer containing 40 mM imidazole and 15 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 0.5 mL/min. After washing, the SpnN protein was eluted with 15 mL of elution buffer containing 100 mM imidazole. The eluted protein was transferred into a dialysis bag and dialyzed against four 1.0 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5) for 1 h per liter of buffer. The dialyzed protein was concentrated by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon) and stored at –80 °C.

Purification of the RfbB Protein. The gene encoding the protein TDP-glucose-4,6-dehydratase, commonly abbreviated as RfbB, from the genome of *E. coli K12* has been cloned into vector pUC18 to give the *rfbB* gene with a 60bp extension containing six successive His codons attached to the 5' end, by Dr. Yun-nan Liu of our laboratory. This construct is subsequently referred to as pYNL-RfbB. Consequently, the expressed protein has a His_6-tag at the *N*-terminus. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli BL21(DE3)-pYNL-RfbB Cells. An overnight culture of *E. coli* BL21(DE3)-pYNL-RfbB, grown in LB medium supplemented with ampicilin (50 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C for another 20 h. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 10 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, pH 8.0), and disrupted with five 1 min sonication bursts, with one min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract was added 10 mL of Ni-NTA slurry (5 mL of resin) and the resulting mixture was incubated for 1.5 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 100 mL of wash buffer containing the same ingredients as the lysis buffer but with 20 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 1.0 mL/min. After washing, the RfbB protein was eluted with 15 mL of the elution buffer (50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, 15% glycerol, pH 8.0). Eluted protein was transferred into a dialysis bag and dialyzed against four 1.0 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5) for 1 h per liter of buffer. The desalted protein was concentrated by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon) and stored at –80 °C. The typical yield was about 120 mg of pure RfbB from a 6 L culture.

Purification of the TylX3 Protein. The gene encoding TylX3 has been cloned into vector pET28b(+) to give the *tylX3* gene with a 60bp extension containing six successive His codons attached to the 5' end, by Dr. Huawei Chen of our laboratory. This construct was referred to as pHC45. The expressed protein has a His_6 tag at the *N*-terminus. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli BL21-pHC45 Cells. An overnight culture of *E. coli* BL21(DE3)-pHC45, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was adjusted to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG at a final concentration of 0.05 mM and allowed to grow for an additional 15 h at 24 °C. The cells were then harvested by centrifugation at 5,000 g for 10 min at 4°C and stored at –80 °C. The typical yield of wet cells was about 8 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with one min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract was added 10 mL of Ni-NTA slurry (5 mL of resin) and the resulting mixture was incubated for 2 h with

gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 10 mL of lysis buffer and various amounts of wash buffers. The wash buffers and elution buffer used contain the same ingredients as the lysis buffer except that the concentrations of imidazole were altered. The column was washed sequentially with 50 mL of wash buffer containing 20 mM imidazole, and 20 mL of wash buffer containing 40 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 1.0 mL/min. After washing, TylX3 protein was eluted with 20 mL of elution buffer containing 250 mM imidazole. Eluted protein was transferred into a dialysis bag and dialyzed against four 1.0 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5) for 1 h per liter of buffer. The dialyzed protein was concentrated by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon) and stored at –80 °C. The typical yield was about 60 mg of pure TylX3 from 6 L culture.

Purification of the TylC1 Protein. The gene encoding TylC1 has been cloned into vector pET24b(+) to give the *tylC1* gene with a 60bp extension containing six successive His codons attached to the 3' end, by Dr. Huawei Chen of our laboratory. This construct was referred to as pHC46. Thus, the expressed protein has a His_6 tag at the *C*-terminus. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli BL21-pHC46 Cells. An overnight culture of *E. coli* BL21(DE3)-pHC46, grown in LB medium supplemented with kanamycin (30 μg/mL) at

37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was adjusted to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG to a final concentration of 0.05 mM and allowed to grow for an additional 15 h at 24 °C. The cells were then harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at – 80 °C. The typical yield of wet cells was about 8 g per liter.

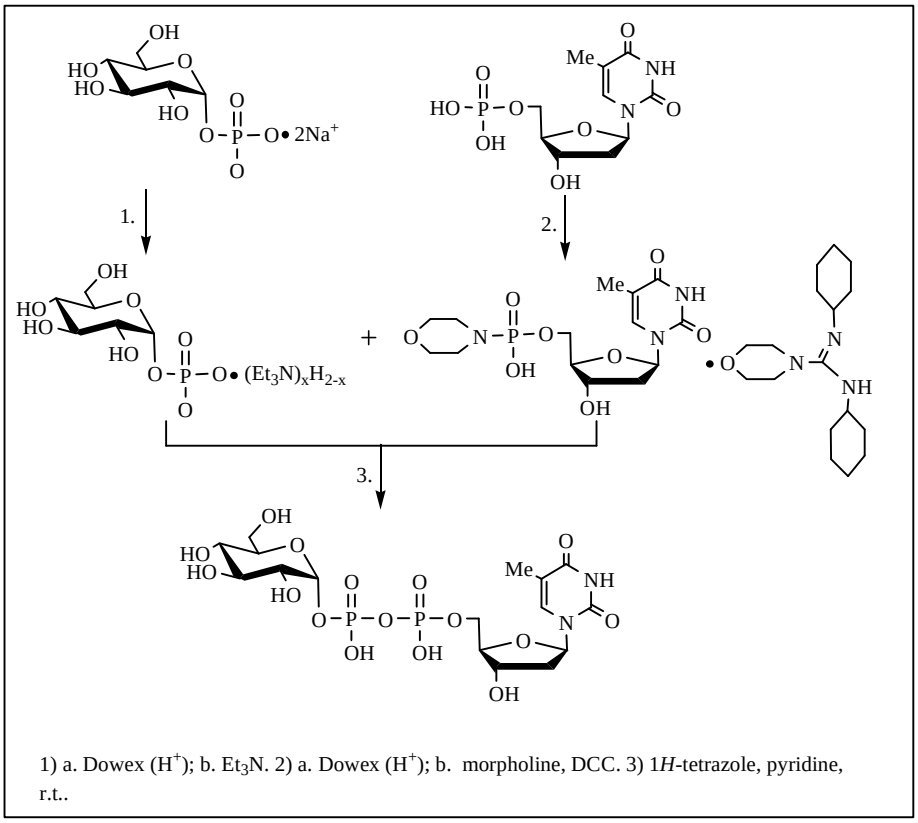
Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with one min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract was added 6 mL of Ni-NTA slurry (3 mL of resin) and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 30 mL of lysis buffer and wash buffers of increasing concentrations of imidazole. The wash buffers and elution buffer used contained the same ingredients as the lysis buffer except that the concentrations of imidazole were altered. The column was washed sequentially with 30 mL of wash buffer

containing 20 mM imidazole, 30 mL of wash buffer containing 40 mM imidazole and 15 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 1.0 mL/min. After washing, the TylC1 protein was eluted with 15 mL of elution buffer containing 250 mM imidazole. The eluted protein was transferred into a dialysis bag and dialyzed against four 1.0 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5) for 1 h per liter of buffer. The dialyzed protein was concentrated by ultrafiltration in an Amicon concentrator using a YM 10 membrane (Amicon) and stored at –80 °C.

Polyacrylamide Gel Electrophoresis. The subunit molecular mass and the purity of the protein samples were assessed by SDS-PAGE. Electrophoresis was carried out in the discontinuous buffer system of Laemmli,[149] and the separating gel and stacking gel were 12% and 4% polyacrylamide, respectively. Prior to electrophoresis, protein samples were mixed with 2 x loading buffer (62.5 mM Tris-HCl, 10% glycerol, 2% SDS, 5% β-mercaptoethanol, and 0.0025% bromophenol blue, pH 6.8) and heated in boiling water for 3 min. Electrophoresis was done in SDS-PAGE running buffer (25 mM Tris-HCl, 192 mM glycine, and 0.1% SDS, pH 8.3) at 200 V. Gel was stained with Coomassie blue and destained with acetic acid/ethanol/water (10: 8: 82 by volume). The subunit molecular mass was estimated by SDS-PAGE as described by Laemmli.[149] Protein standards for SDS-PAGE include α-lactalbumin (14 kDa), trypsin inhibitor (20 kDa), trypsinogen (24 kDa), carbonic anhydrase (29 kDa), glycerol-3-phosphate dehydrogenase (36 kDa), egg albumin (45 kDa), and bovine serum albumin (66 kDa).

Molecular Weight Determination. Native molecular weights of SpnO and SpnN were measured by size exclusion chromatography using a Pharmacia FPLC equipped with a Superdex 200 HR 10/30 column. The buffer (50 mM sodium phosphate, 0.15 M NaCl, pH 7.0) was recommended by the column manufacturer, A flow rate of 0.5 mL/ min was used in the analysis. Calibration of the column was achieved using the following protein standards (Sigma): cytochrome c (12.4 kDa), carbonic anhydrase (29



Scheme 2-4 . Synthetic scheme for TDP-glucose.

kDa), bovine serum albumin (66 kDa), alcohol dehydrogenase (150 kDa), and β-amylase (200 kDa). The void volume (Vo) of the column was measured using blue dextran. A linear fit to a plot of the molecular weight versus V_e/V_o was used to estimate the native molecular mass (M_r).[151]

Synthesis of TDP-glucose. Chemical synthesis of TDP-glucose was carried out according to the scheme shown in Scheme 2-4.*Preparation of TMP (H^+).* Dowex cation exchange resin was packed in a glass column (2.5 x 22 cm) and was charged with hydrogen cation by washing with 1 L of 8% (w/v) HCl. Excess acid was removed by extensive washing with H_2O until the pH of the eluent was between 5 and 6. TMP disodium salt was dissolved in 2 mL of H_2O and applied to the top of the resin. The column was washed with water by gravity. The fraction size of the eluent was set at 8 mL. The fractions whose pH values were near 1 as tested by pH paper were combined and subjected to lyophilization.

Preparation of 1H-tetrazole. The commercial 1*H*-tetrazole is a 3% (w/v) solution in acetonitrile. Evaporation *in vacuo* gave the solid white powder 1*H*-tetrazole. The identity of the compound was confirmed by ^{1}H NMR. ^{1}H NMR (D_2O): δ 9.00 (1H, s).

Preparation of TMP morpholidate. The synthesis was carried out according to the procedure described by Moffatt and Khorana.[152] A solution of dicyclohexyl-carbodiimide (4.126 g, 20 mmol) in *t*-butyl alcohol (45 mL) was added dropwise to a refluxing solution of TMP (5 mmol of free acid) in a mixture of water (30 mL), *t*-butyl alcohol (30 mL) and 99.5% morpholine (1.7 mL). The addition was completed in 10 h, and the mixture was further refluxed overnight. When the mixture was cooled down to

room temperature, any crystalline material present was removed by filtration and washed with *t*-butyl alcohol. The filtrate was evaporated *in vacuo* until *t*-butyl alcohol was largely removed. The remaining aqueous phase was extracted three times with ether. The clear aqueous solution was then evaporated to dryness *in vacuo* and the last traces of water were removed by lyophilyzation. The glassy residue was transferred as a solution in a minimum volume of methyl alcohol to three 50-mL centrifuge tubes and the volume of the methyl alcohol in each centrifuge tube was carefully reduced *in vacuo* to about 4 mL. The addition of dry ether (35 mL) precipitated a gummy solid which on precipitation with fresh ether changed to dry white powder. After a further wash with dry ether, the product was dried *in vacuo* at room temperature. The TMP-morpholidate was obtained as salt of 4-morpholine *N*,*N*'-dicyclohexylcarboxamidine in 90% yield. ^{31}P NMR (D_2O): δ 8.42 ppm. Low-resolution ESI-MS calculated for $C_{30}H_{53}N_6O_8P$ $(M-H)^-$ was 655, and a mass of 655 was found.

Preparation of Glucose-1-phosphate Triethylamine. Dowex cation exchange resin was packed in a glass column (2.5 x 22 cm) and charged with hydrogen cation by washing with 1 L of 8% (w/v) HCl. Excess acid was removed by extensive washing with H_2O until the pH of the eluent was between 5 and 6. D-glucose-1-phosphate sodium salt (1.2 g) was dissolved in 5 mL of H_2O and applied to the top of the resin. The column was washed with H_2O and the eluent was collected in 8 mL fractions. The fractions whose pH values were near 1 judged by pH paper were combined, and 2 mL of triethylamine was added. After mixing well, the sample was dried by lyophilization.

Preparation of TDP-glucose. Glucose-1-phosphate triethylamine salt (540 mg, 1.5 mmol), and TMP-morpholidate (1.54 g, 2.25 mmol) were dissolved in 8 mL of anhydrous pyridine and the solution was evaporated to dryness *in vacuo*. The process of dissolving and evaporation was repeated twice with 1*H*-tetrazole (190 mg, 2.7 mmol) added to the solution the second time. After the last trace amounts of water had been removed, the mixture was dissolved in 12 mL of anhydrous pyridine and stirred at room temperature under nitrogen for 3 days. The solvent was removed by evaporation *in vacuo*, while the residue was resuspended in 20 mL of water and extracted with chloroform (3 x 15 mL). The aqueous portion was dried by lyophilization. The dried compounds were redissolved in 2 mL of water and applied to Bio-Rad P2 (extra fine) column (2.5 x 120 cm). The column was run with 25 mM NH_4HCO_3 buffer at a flow rate of 0.24 mL/min. The fraction size was 6 mL. The absorbance of each fraction at 267 nm was examined with a spectrophotometer and the ones that showed strong absorbance were lyophilized as individual fractions. The purity of each fraction was inspected by ^{1}H NMR. The fractions that contained pure TDP-glucose were saved, while the ones with impurities were further purified with Dowex cation exchange column. The impure fractions was concentrated to 12 mg/mL and applied to the Dowex cation exchange column (2.5 x 20 cm). The compounds were eluted with water with a fraction size between 2 and 3 mL. The fractions collected from the Dowex column whose pH values were near 1 as judged by pH paper were neutralized with ammonium bicarbonate and lyophilized individually. The purity the TDP-glucose obtained was inspected by ^{1}H NMR. ^{1}H NMR (D_2O): δ 1.78 (3H, s, 5’’-Me), 2.22 (2H, m, 2’-H), 3.28-3.40 (2H, m, 2-

H, 3-H), 3.59-3.78 (2H, m, 4-H, 5-H), 3.95-4.06 (3H, m, 4'-H, 5'-H), 4.45 (1H, m, 3'-H), 5.44 (1H, dd, *J* = 6.9, 3.3 Hz, 1-H), 6.20 (1H, t, *J* = 6.9 Hz, 1'-H), 7.59 (1H, s, 6''-H). ^{31}P NMR (D_2O): δ –10.30 (d, *J* = 21.4 Hz), -11.83 (d, *J* = 21.4 Hz).

Preparation of TDP-4-keto-6-deoxy-D-glucose (the RfbB product). This compound was prepared by the RfbB enzymatic reaction with TDP-glucose as the substrate. The 1.5 mL reaction mixture contained 30 mg of TDP-glucose and 1.5 mg of RfbB in 20 mM Tris-HCl buffer (pH 7.5). The reaction was allowed to proceed for 2 h at 37 °C. An HPLC based assay with an Adsorbsphere SAX analytical column (5 u, 4.5 x 250mm) was used to monitor the progress of the reaction. A linear gradient from 20% to 60% buffer B (500 mM potassium phosphate, pH 3.5) in buffer A (50 mM potassium phosphate, pH 3.5) over 20 min gave satisfactory separation between the substrate (retention time = 12.5 min) and the product (retention time =13.2 min). The RfbB protein was removed with Centricon YM-10 microconcentrator (Amicon), and the filtrate was used directly without further purification. Spectra of RfbB product: ^{1}H NMR (D_2O): δ 1.04 (3H, d, *J* = 6.6 Hz, 5-Me), 1.75 (3H, s, 5''-Me), 2.14-2.24 (2H, m, 2'-Hs), 3.44 (1H, ddd, *J* = 10.0, 3.6, 3.3 Hz, 2-H), 3.60 (1H, d, *J* = 10.0 Hz, 3-H), 3.92 (1H, q, *J* =6.6 Hz, 5-H), 3.95-4.06 (3H, m, 4'-H, 5'-Hs), 4.40-4.48 (1H, m, 3'-H), 5.36 (1H, dd, *J* = 7.0, 3.6 Hz, 1-H), 6.17 (1H, t, *J* = 6.9 Hz, 1'-H), 7.56 (1H, s, 6''-H).

Activity Assays for SpnN.

a. UV-absorption Activity Assay. The activity of SpnN was examined spectrophotometrically by following the consumption of NADPH at 340 nm (ε_{340} = 6220

M^{-1} cm^{-1}). A typical assay mixture of 100 μL contained 0.85 mM of the RfbB product, 0.18 mM of NADPH, 2 μg of TylX3 and 4 μg of SpnN in 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of TylX3 and was carried out at room temperature (24 °C). The absorbance decrease at 340 nm was linear for the first 5 min.

b. HPLC Activity Assay. The HPLC assay was developed for SpnN by monitoring the formation of new product(s) at 267 nm with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) connected to a CarboPac™ PA1 guard column (4 x 50 mm). A typical assay mixture of 50 μL contained 0.42 mM of TDP-Glu, 80 μg of RfbB, 0.1 mM of NADPH, 1 μg of TylX3, 2 μg of of SpnN, and 50 mM potassium phosphate (pH 8.0). The reaction was initiated by the addition of TylX3 and was carried out at room temperature (24 °C). A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by dilution with a 10-fold excess of water and flash freezing in liquid nitrogen. The enzyme was removed by filtering through Centricon YM-10 microconcentrator by centrifugation. The filtrate was then subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60%B over 20 min was used to resolve the substrate TDP-glucose (retention time = 41.5 min) and the SpnN product (retention time = 33.4 min). Buffer A was water and Buffer B was 500 mM NH_4OAc, pH 7.0, adjusted with diluted $NH_3 \cdot H_2O$. The elution gradient was followed by a final wash with 100% B for 5 min, and a reequilibration with 5% B for 15 min.

Determination of the Reducing Agent. Two reducing agents, NADH and NADPH were examined for their ability to reduce the SpnO product in the SpnN catalyzed reaction. The compounds were individually added to a final concentration of 0.15 mM in the SpnN UV-absorption activity assays described above. The efficiency of the two as reducing agents was compared by calculating the initial rate of absorbance decrease at 340 nm.

Catalytic Activity of SpnN Compared with TylC1. Two enzyme pairs, TylX3/SpnN and TylX3/TylC1 were examined for their efficiency to produce each corresponding 2-deoxygenated sugar. The UV-absorption assays were carried out as described before with 0.92 μM of TylX3 and 2.0 μM of SpnN or TylC1 in the reaction mixture. The catalytic activities of the two enzyme pairs were compared by calculating the initial rate of absorbance decrease at 340 nm.

Activity Assays for SpnO.

a. UV-absorption Activity Assay. The activity of SpnO was determined spectrophotometrically by following the consumption of NADPH at 340 nm (ε_{340} = 6220 M^{-1} cm^{-1}). A typical assay mixture of 100 μL contained 17 μM of TDP-6-deoxy-4-keto-glucose, 18 μM of NADPH, 0.4 μg of SpnN and 0.2 μg of SpnO in 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of SpnO and was carried out at room temperature (24 °C).

b. HPLC Activity Assay. The HPLC assay that was developed for SpnO was carried out by monitoring the formation of the degradation product, maltol, at 267 nm with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) connected to a

CarboPac^{TM} PA1 guard column (4 x 50 mm). A typical assay mixture of 50 μL contained 30 μM of RfbB product, 1 mM of $MgCl_2$, 4.2 μM of SpnO, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of SpnO and was carried out at room temperature (24 °C). A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by a 10-fold dilution in water and flash freezing in liquid nitrogen. The enzyme was removed by filtering through a Centricon YM-10 microconcentrator with centrifugation. The filtrate was then subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60%B over 20 min was used to resolve the substrate RfbB product (retention time = 35.3 min) and the degradation product (retention time = 2.4 min). Buffer A was water and Buffer B was 500 mM NH_4OAc, pH 7.0 adjusted with diluted $NH_3 \cdot H_2O$. The elution gradient was followed by a final wash with 100% B for 5 min, and a reequilibration with 5% B for 15 min.

Isolation and Characterization of the SpnN product. A preparative incubation contained 25 mg of TDP-glucose (41.8 μmol), 2.1 mg of RfbB, and 21.6 mg of NADPH (26.0 μmol) in 1.5 mL of 50 mM Tris-HCl buffer (pH 7.5, containing 10% glycerol). The reaction was initiated by the addition of TylX3 (15 nmol) and SpnN (25 nmol). Due to the instability of the TylX3 product, an excess of SpnN was used. The reaction was incubated at room temperature and was followed by monitoring the consumption of NADPH at 340 nm. Usually, the decrease of the absorbance at 340 nm stopped after 4 h. The enzymes were removed with a Centricon YM-10 microconcentrator (Amicon), and the filtrate was then separated by a gel filtration column (Bio-Rad P2, extra fine, 2.5 x

120 cm) that had been preequilibrated with 25 mM NH_4HCO_3. The column was run with the same buffer at a flow rate of 12 mL/h at 4 °C. Fractions were analyzed spectrophotometrically at 267 nm and 340 nm. The fractions that displayed a single absorption maximum at 267 nm were lyophilized individually and analyzed by NMR. Spectra data for the TylX3/SpnN product: ^{1}H NMR (D_2O): δ 1.07 (3H, d, *J* = 6.3 Hz, 5-Me), 1.71 (1H, m, 2-H_{ax}), 1.79 (3H, s, 5''-Me), 2.01 (1H, ddd, *J* = 13.2, 5.2, 1.2, 2-H_{eq}), 2.22 (2H, m, 2'-H), 3.88 (1 H, dd, *J* = 11.7, 5.1 Hz, 3-H), 3.90 (1H, q, *J* = 6.6 Hz, 5-H), 4.03 (3H, m, 4'-H, 5'-H), 4.48 (1H, m, 3'-H), 5.48 (1H, dd, *J* = 6.9, 3.3 Hz, 1-H), 6.21 (1H, t, *J* = 7.2 Hz, 1'-H), 7.62 (1H, s, 6''-H). ^{13}C NMR (D_2O): δ 11.6, 11.9, 36.1 (d, *J* = 6.3 Hz), 38.9, 68.3, 70.3, 71.2, 85.2, 85.6 (d, *J* = 9.3 Hz), 93.3, 94.7 (d, *J* = 5.5 Hz), 112.0, 137.6, 152.0, 166.9, 207.2. ^{31}P NMR (D_2O): δ –10.7 (d, *J* = 21.2 Hz), -12.7 (d, *J* = 21.2 Hz). High-resolution FAB-MS calculated for $C_{16}H_{23}N_2O_{14}P_2$ $(M-H)^-$ was 529.0624, and a mass of 529.0627 was found.

3. RESULTS AND DISCUSSION

Cloning and Overexpression of the SpnO Protein. The protein encoded by *spnO* gene has significant sequence similarity to several proteins that are proposed to be dehydratases in deoxysugar biosynthesis. In order to clearly demonstrate the role that SpnO plays in TDP-D-forosamine biosynthesis, we decided to isolate the protein. Three clones were made with *spnO* gene ligated to either the pET28b(+) or the pET24b(+) vector. Both pET24b and pET28b carry a T7 promoter. A high level of expression of T7

RNA polymerase is induced in the BL21 host strain by addition of IPTG, thereby inducing a high level of expression of the gene downstream of the promoter.

The *N*-terminal His_6-tagged SpnO was highly expressed. When the whole cell lysate of the pLHON-transformed cells was solubilized with SDS and analyzed by SDS-PAGE, there was a significant enrichment of a protein of molecular weight of 56 kDa, the size of SpnO, as compared to the negative control. In the negative control, the expression cells were transformed with empty pET28b(+) vector and were grown under the same conditions as the expression cells transformed with the pLHON clone. The natural gene without any tag when expressed in *E. coli* induced with IPTG, produced a moderate amount of the 54 kDa protein as judged by SDS-PAGE analysis of the solubilized whole cell lysate. However, a solubility test, in which only soluble proteins were analyzed by SDS-PAGE, showed that neither the *N*-terminal tagged SpnO nor the native SpnO was soluble to a significant extent. Most of the expressed protein formed inclusion bodies and precipitated after the cells were lysed. The *C*-terminal His_6-tagged SpnO was poorly expressed, and thus, the protein was not subjected to a solubility test. The cells transformed with pLHON were used for further expression tests.

Various expression conditions were examined. Expression at 15 °C did not significantly improve the protein solubility. On the other hand, the cell yield of the culture grown at 15 °C was much lower than that grown at higher temperatures. The expression condition found to be optimal was to grow the culture at 24 °C for 20 h after IPTG induction.

Purification and Characterization of the SpnO Protein (Figure 2-3). *N*-terminal His_6-tagged SpnO protein was purified to near homogeneity by a single Ni-NTA affinity chromatographic step. Precipitation of the protein during dialysis was observed. Inclusion of 20% glycerol was helpful in stabilizing the protein, but precipitation could not be completely avoided. The yield of SpnO from a 6 L culture, which was 2 mg, was much lower than the 60 mg obtained from purification of the homologous protein, TylX3. The subunit molecular weight of 56 kDa as estimated by SDS-PAGE correlated well with the predicted value of 56,553 Da calculated from the deduced amino acid sequence (plus the *N*-terminal His_6-tag). A M*r* of 96 kDa, as judged by gel filtration chromatography, indicates that the native SpnO exists as a homodimer. The standard curve for molecular weight determination is shown in Figure 2-4. The absorption spectrum of the purified protein shows no absorbance above 300 nm.

Cloning and Expression of the SpnN Protein. To examine the function of the *spnN*-encoded enzyme, the gene was cloned into various expression vectors. Numerous attempts to obtain a usable yield of soluble SpnN were carried out by Dr. Zhao. Originally, the start codon of the gene was thought to exist eight nucleotides downstream of the *spn*O gene (Figure 2-5). This gene, named *spnN*-short, was ligated into several pET vectors. However, none of these clones gave a decent expression of *spnN*-short. Several strategies, including ligation of *spnN*-short into pET vectors at different

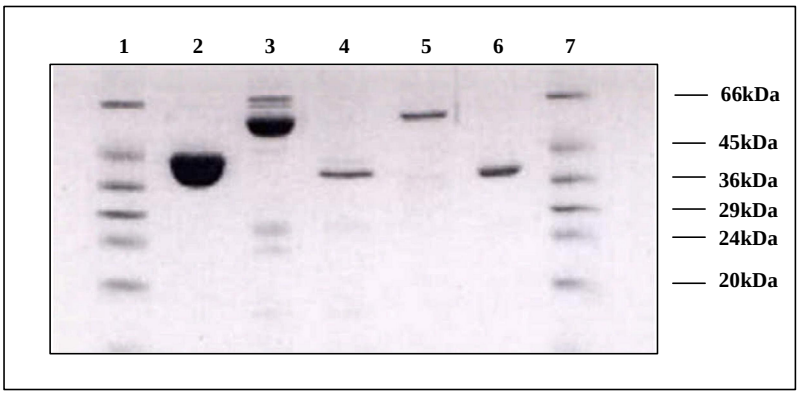


Figure 2-3. SDS-polyacrylamide gel electrophoresis of proteins obtained in Chapter 2. Lanes 1 and 7, molecular markers; lane 2, RfbB protein; lane 3, TylX3 protein; lane 4, TylC1 protein; lane 5, SpnO protien; lane 6, SpnN protein.

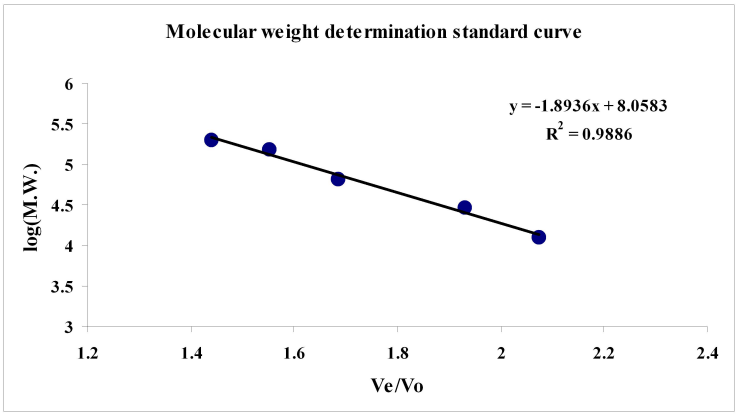


Figure 2-4. Molecular weigh determination standard curve.

restriction sites, and insertion of an *E. coli* ribosome binding site (RBS) directly upstream of *spnN*-short were tried, with no improvement in the protein expression observed. In addition, various expression hosts were tested, such as *E. coli* BL21(DE3), BL21(DE3)

pLysS, HMS174(DE3), Rosetta^{TM}2(DE3), RosettaBlue-^{TM}2(DE3), BL21-CodonPlus(DE3)-RP, and yeast *Pichia pastoris*. Again, no expression of *spnN*-short was observed. Later, close inspection of the gene cluster revealed that there was another ATG start codon which overlapped with the stop codon of the *spnO* gene (Figure 2-5, a.). This ATG codon is in frame with *spnN*-short, and begins 12 nucleotides upstream of the *spnN*-short start codon. Translation initiation at this upstream start codon would result in

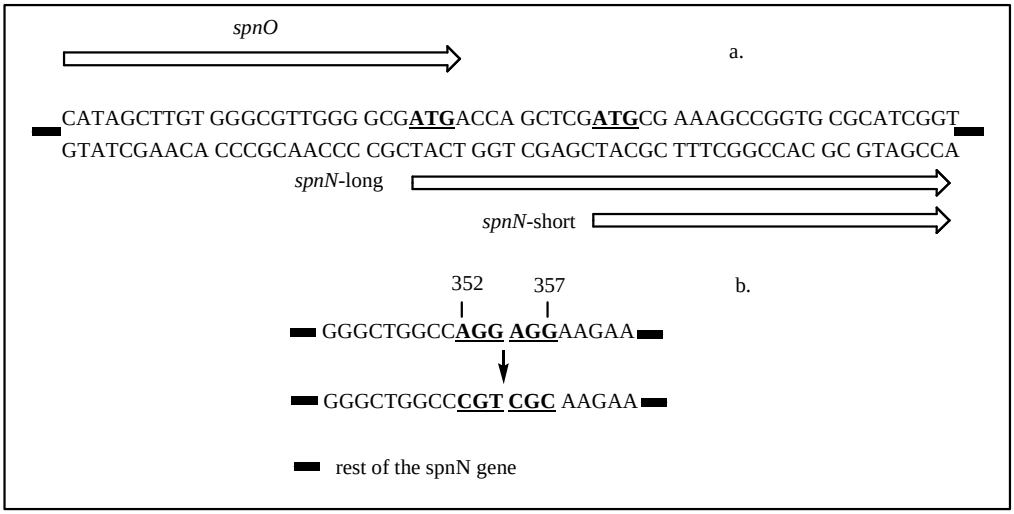


Figure 2-5. (a). The *spn*N and *spn*O gene arrangement. There are two possilbe ORFs for *spn*N: one starts 8 nts downstream of *spn*O gene; the other one starts just before the end of *spn*O gene. (b). Mutation of rare Arg codons in *spn*N to commonly used codons in *E.coli*. The numbers indicate the location of the rare codons in *spn*N-long gene.

addition of 4 amino acids, MTSS, to the *N*-terminus of SpnN-short. This open reading frame (ORF), named *spnN*-long, was cloned into pET24b(+) vector. The resulting vector was expressed in BL21(DE3), and showed a small improvement in the yield of SpnN protein. However, a usable amount of SpnN could not be isolated from a large-scale preparation using this clone. The sequence of *spnN*-long was scrutinized again. It was

found that there is one occurrence of AGGAGG within the gene sequence, occurring from nucleotide 352 to nucleotide 357 (Figure 2-5, b.). AGG encodes Arg, but is an extremely rare codon in the *E. coli* genome. The occurrence of two rare codons in tandem could make expression of the gene very difficult. In an attempt to overcome this problem, the AGGAGG sequence was mutated to CGTCGC. Both CGT and CGC are more commonly used codons encoding Arg in *E. coli*. The clone containing the optimized *spnN*-long gene was transformed into *E. coli* BL21-CodonPlus(DE3)-RP expression strain. The SpnN protein was finally expressed in high quantity after IPTG induction, as judged by SDS-PAGE and Western-Blot of the whole cell lysate which had been solubilized by SDS.

Purification and Characterization of the SpnN Protein (Figure 2-3). The protein was purified to near homogeneity by a single Ni-NTA chromatographic step. However, the binding between SpnN and the Ni-NTA resin was not tight. The protein began to elute from the resin along with impurities when the imidazole concentration in the washing buffer rose to 40 mM. The remaining portion of SpnN that eluted with 100 mM imidazole was nearly pure. About 4 mg of SpnN protein was obtained from 6 L of cell culture. Considering the high level of expression of the protein, such a low yield indicated that most of the protein that was expressed in the host cells formed inclusion bodies. Expression at lower temperatures, such as 15°C, did not improve the solubility of the protein. The subunit molecular weight of 37 kDa estimated by SDS-PAGE is consistent with a molecular weight 37602 Da deduced from the translated amino acid sequence of SpnN (including the His_6-tag). The molecular weight of native SpnN was

estimated to be 38 kDa by gel filtration chromatography, suggesting that the protein is a monomer in solution. The absorption spectrum of the purified protein shows no absorbance above 300 nm.

Synthesis of TDP-glucose. Because TDP-glucose is the starting material in the TDP-D-forosamine biosynthetic pathway, a large amount of TDP-glucose is required in order to isolate the enzymatic products formed in the later steps in the forosamine biosynthesis. The inhibitory cost of TDP-glucose from commercial sources prompted us to resort to chemical synthesis. The commercially available starting materials TMP and glucose-1-phosphate were both in sodium salt form, making them less soluble in organic solvent and inactive for organic reactions. By addition of the sodium salts of these compounds to a Dowex cation exchange column that had been charged with H^+, glucose-1-phophate and TMP were obtained in their acidic (H^+) forms. To improve the yield of the coupling reaction with TMP-morpholidate, the acidic form of glucose-1-phophate was further reacted with triethylamine to form glucose-1-phosphate triethylamine salt, which has improved solubility in pyridine. The synthesis of TMP-morpholidate was done according to the procedure established by Moffatt and Khorana (90% yield).[152] The compound was stable enough to tolerate overnight reflux. The coupling reaction between glucose-1-phosphate triethylamine and TMP-morpholidate requires an anhydrous environment. Trace amounts of water in the solvent reduce the product yield. The reaction was monitored by ^{31}P NMR and was quenched when the doublet peaks at around –10 ppm and –12 ppm stopped increasing. Half of the TDP-glucose formed was isolated in pure form by Bio-Gel P-2 chromatography. The remainder was further purified by

Dowex cation exchange chromatography. The total yield of TDP-glucose was about 80%.

Purification of the TylX3 and RfbB Proteins (Figure 2-3). TDP-6-deoxy-4-keto-glucose is a common intermediate in TDP-deoxysugar biosynthetic pathways. However, the enzyme catalyzing the conversion of TDP-glucose (**2**) to TDP-6-deoxy-4-keto-glucose (**3**) could not be identified in the TDP-forosamine biosynthetic gene cluster. *In vivo*, TDP-6-deoxy-4-keto-glucose (**3**) used for TDP-D-forosamine biosynthesis is likely obtained from general cellular pools, since the compound is used as a precursor in other pathways, known to exist in Actinomycetes.[147,165,166] To characterize the enzymatic steps involved in TDP-D-forosamine biosynthesis *in vitro,* an enzyme that could convert TDP-glucose (**2**) to TDP-6-deoxy-4-keto-glucose (**3**) is needed. Thus, the RfbB protein, a TDP-glucose-4,6-dehydratase from *E. coli K12* was purified. An *N*-terminal His_6-tag greatly facilitated RfbB purification procedures and the yield of the soluble protein was quite high. The protein could be purified to near homogeneity by a single Ni-NTA chromatographic step. About 120 mg of pure RfbB protein was obtained from a 6 L culture grown in LB medium.

Since the yield of the purified SpnO protein was low and the protein showed significant instability, an SpnO replacement was needed in order to isolate the product from the SpnO/SpnN coupled reaction. TylX3, the protein that catalyzes 2,3-dehydration of TDP-6-deoxy-4-keto-glucose (**3**) in the TDP-mycarose biosynthetic pathway, shows 43.8% sequence identity to the SpnO protein and so was chosen to replace SpnO. The *N*-terminal His_6-tagged protein was purified to near homogeneity by a single Ni-NTA

chromatographic step. About 60 mg of pure protein was obtained from a 6 L culture grown in LB medium. An amount of 15% glycerol was included in all buffers to reduce protein precipitation during purification. Although some precipitation occurred, a significant amount of TylX3 remained stable. It is interesting to observe that the two homologous proteins, TylX3 and SpnO, displayed different solubilities in the expression host cells, and different binding affinities to the Ni-NTA resin. These differences lead to different yields of the proteins.

Purification of the TylC1 Protein (Figure 2-3). In the TDP-L-mycarose biosynthetic pathway (Scheme 2-3), TylC1 catalyzes the C-3 ketoreduction of the TylX3 product, a function analogous to that proposed for SpnN, except that the stereochemistry of the C-3 hydroxyl group is predicted to be equatorial in the SpnN reaction, as opposed to axial, as was observed in previous studies by our group in the TylX3/TylC1 reaction.[71] It would be interesting to compare the catalytic rates of the two enzymes SpnN and TylC1. Thus, the *C*-terminal His_6-tagged TylC1 protein was purified to near homogeneity by a single Ni-NTA chromatographic step. About 4 mg of pure TylC1 was obtained from a 6 L culture grown in LB medium. The inclusion of 15% glycerol in all the buffers was crucial to prevent TylC1 from precipitating during the purification.

Activity Assays for SpnN. TylX3 has been characterized as the hexose 2,3-dehydratase in the TDP-L-mycarose biosynthetic pathway[71]. The TylX3 product was previously shown to be reduced by the NADPH-dependent catalysis of the reductase TylC1. Accordingly, in the UV activity assay for SpnN, the reduction of the TylX3 product was quantified by following the decrease of A_{340} caused by NADPH oxidation.

In the HPLC assay, the disappearance of TDP-glucose and the simultaneous formation of the SpnN product were observed.

Reducing Agent for SpnN. NADPH has been shown to be the substrate in the Gra Orf26 and TylC1 catalyzed reduction reactions. On the other hand, sequence analysis demonstrated that a number of ketoreductases and oxidoreductases that have significant sequence similarity to SpnN use NAD(H) as cosubstrate or cofactor. Both NADH and NADPH were examined as possible reducing agents in the SpnN-catalyzed reaction. It was found that NADH and NADPH could each act as a reducing agent in the SpnN-catalyzed reaction, although NADPH was preferred (Figure 2-6). The rate difference of the reactions using these two reducing agents was about 2 folds.

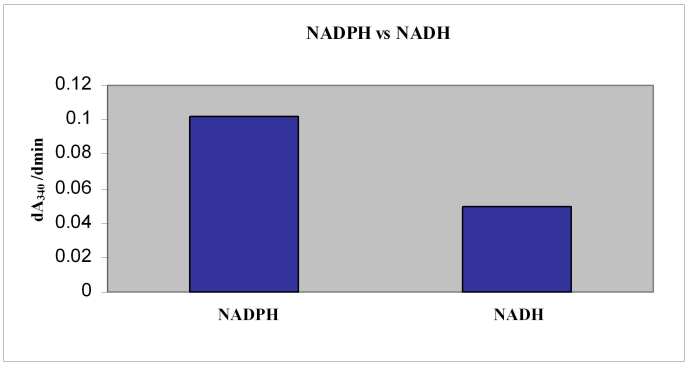


Figure 2-6. Reducing agent determination: NADPH vs. NADH.

Catalytic Activity of SpnN vs. TylC1. Since both enzymes pairs, SpnO/SpnN and TylX3/TylC1, catalyze C-2 deoxygenation in deoxysugar biosynthesis, it would be interesting to compare the catalytic efficiency of the two pairs. However, the fact that SpnO could only demonstrate its *in vitro* catalytic activity at low substrate concentrations

precluded accurate measurement of the activity of the SpnO/SpnN pair. Nonetheless, it would still be informative to compare the catalytic rates of TylC1 and SpnN. In TylX3/TylC1 and TylX3/SpnN coupled assays using the same concentration of TylX3, SpnN was demonstrated to be a superior reductase compared with TylC1, as shown by a 20-fold increase in NADPH consumption for the TylX3/SpnN pair over the TylX3/TylC1 pair (Figure 2-7). This result is intriguing since TylX3/TylC1 is a natural enzyme pair, participating in TDP-L-mycarose biosynthesis, while TylX3 and SpnN are from different deoxysugar biosynthetic pathways. It is possible that the inferior catalytic activity demonstrated by TylX3/TylC1 might be due to the stereochemistry of the hydroxyl group in the product. The axial hydroxyl group of the TylC1 product (**28**, Scheme 2-3) is thermodynamically unstable compared with the equatorial hydroxyl group of the SpnN product (**24**, Scheme 2-1) and thus might be more difficult to form.

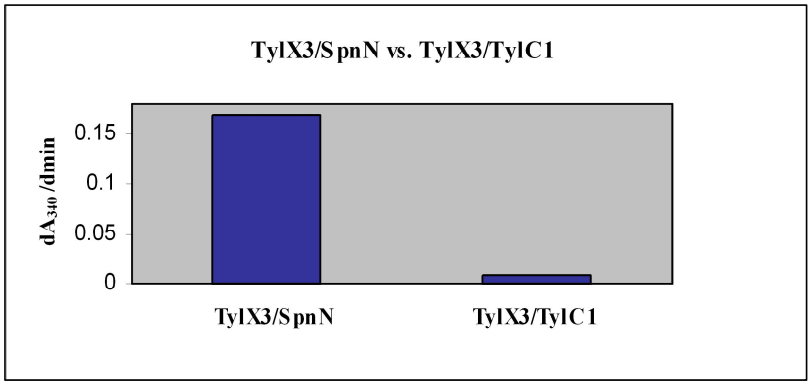


Figure 2-7. Comparison of reduction efficiency between SpnN and TylC1.

Activity Assays for SpnO. SpnO has displayed low solubility in expression cells and poor binding affinity to Ni-NTA resin, which is quite different from the previously characterized TylX3. However, a more unexpected difference came from the catalytic activities that the two enzymes displayed. TylX3 has been demonstrated to be a quite efficient enzyme.[129] When the enzyme was mixed with TDP-6-deoxy-4-keto-glucose (**3**, Scheme 2-3), the degradation product maltol (**66**, Scheme 2-3) quickly formed in the absence of a reductase.[70,71,129] Moreover, in the UV activity assay for SpnN, the decrease of A_{340} was obvious after the addition of TylX3, revealing the efficiency of the TylX3/SpnN pair. In contrast, the UV activity assay for SpnO showed that A_{340} did not decrease at all when the concentration of the substrate TDP-6-deoxy-4-keto-glucose was higher than 0.2 mM. Since SpnN had been proved to be an active enzyme in the activity assays described previously, the failure to observe the absorbance decrease at 340 nm in the SpnO/SpnN coupled assay was ascribed to SpnO. Originally, it was speculated that the presence of the *N*-terminal His_6-tag in the recombinant protein interfered with the enzymatic activity. However, removing the His_6-tag by thrombin cleavage did not yield a more active protein. Only when the concentration of TDP-6-deoxy-4-keto-glucose was decreased to lower than 0.2 mM, was a slight decrease of A_{340} observed (Figure 2-8, a). Lower concentrations of TDP-6-deoxy-4-keto-glucose were tested in an attempt to improve SpnO catalytic activity. The extent of A_{340} decrease was only slightly greater when the substrate concentration was reduced to 15 μM compared with that observed at 0.2 mM substrate concentration. Mg^{2+} did not improve the efficiency of SpnO.

SpnO was thought to catalyze the same reaction as TylX3 in the forosamine biosynthetic pathway. The authentic TylX3 product undergoes degradation to yield TDP and maltol (**66**) in the absence of TylC1. Thus, when incubating TDP-6-deoxy-4-keto-

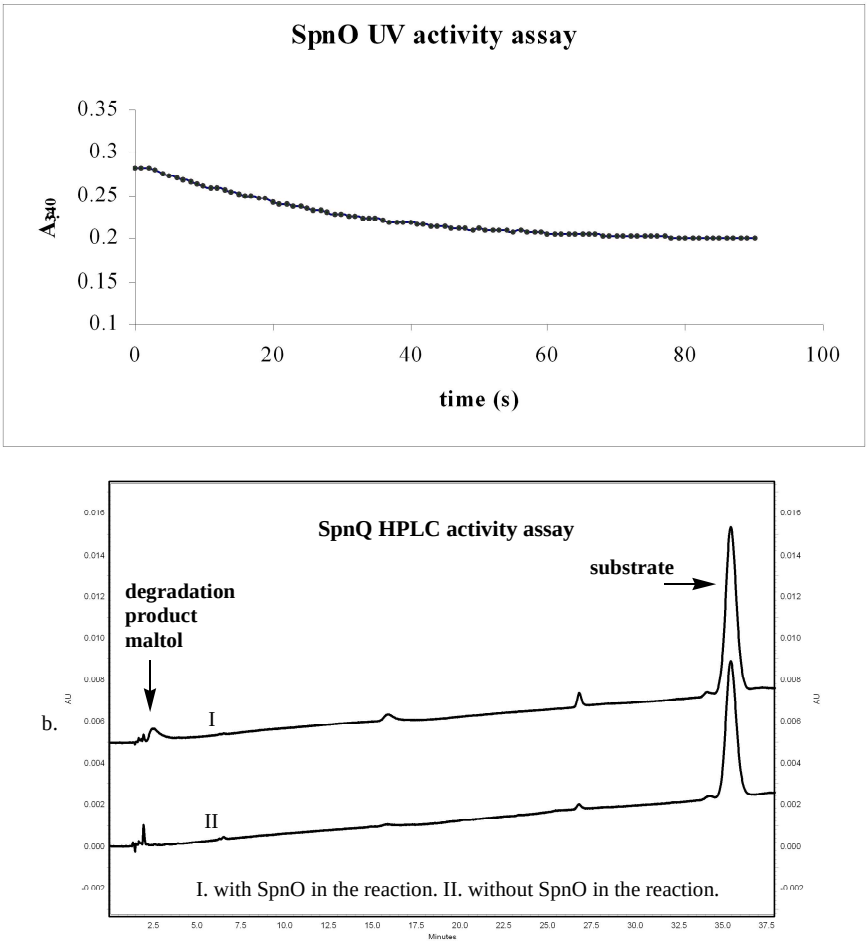


Figure 2-8. SpnO activity assays. (a). UV-absorption assay. (b). HPLC assay. The assay conditions were described in the Experimental Procedures section.

glucose with SpnO, the products are expected to be TDP and maltol. Formation of maltol was immediately observed after the addition of SpnO when the reaction was monitored

by HPLC (Figure 2-8, b). Since the substrate TDP-6-deoxy-4-keto-glucose contained a small amount of TDP, a slight increase in the amount of TDP was difficult to confirm. Although the size of the maltol peak in HPLC traces of the SpnO reaction is small, its presence was definite by comparison of the trace with that of a control reaction in which SpnO was omitted. This is also consistent with the minor A_{340} decrease observed in the SpnO reaction when monitored by UV. Consistent with the UV activity assay, the degradation products of the SpnO reaction could only be observed when the substrate concentration was lower than 0.2 mM. In spite of the low conversion percentage, the formation of maltol was quick and reached near maximum level within 1 min. Extending the incubation time to 1 h only slightly increased the amount of maltol formed.

In summary, SpnO can catalyze the 2,3-dehydration of the substrate TDP-6-deoxy-4-keto-glucose (**3**) to form TDP-2,6-dideoxy-3,4-diketo glucose (**23**), which in the presence of SpnN and NADPH was reduced to TDP-2,6-dideoxy-4-keto-glucose. In the absence of SpnN and NADPH, the SpnO product was hydrolyzed to form TDP and maltol (**66**). However, SpnO was catalytically less competent than the homologous TylX3. One justification for these results is that the conditions of an *in vitro* assay where product accumulates, are quite different from the *in vivo* system where the SpnO catalyzed reaction is coupled to both upstream and downstream reactions. *In vivo*, the substrate is produced gradually by upstream pathway enzymes, and the product of the SpnO reaction may be quickly channeled to the downstream enzymes in the TDP-D-forosamine biosynthetic pathway allowing SpnO to achieve its maximal catalytic rate without encountering substrate or product inhibition. It is also possible that the

production of TDP-L-rhamnose and TDP-D-forosamine are fine-tuned by the concentration of TDP-6-deoxy-4-keto-glucose by the action of SpnO. When the concentration is high, 2,3-dehydration of TDP-6-deoxy-4-keto-glucose may be minor compared to 3,5-epimerization of it to initiate TDP-L-rhamnose biosynthesis. SpnO may have evolved to be less competent so that the majority of TDP-6-deoxy-4-keto-glucose in the general cellular pool is routed toward production of other products besides TDP-D-forosamine. A full understanding of the differences that TylX3 and SpnO have displayed must await further investigation. Study of the interaction between SpnO and SpnN to support possible substrate channeling of the unstable SpnO product using two-hybrid system in Yeast is ongoing.[153]

Isolation and Characterization of the TylX3/SpnN Product. The SpnN product was isolated from a one-pot reaction mixture containing TDP-glucose, RfbB, NADPH, TylX3 and SpnN. The use of TylX3 instead of SpnO in the reaction was due to the low yield and the instability of the latter protein. It had been explained earlier that TylX3 and SpnO share 43.8% identity and catalyze identical reactions. Thus, the replacement of SpnO with TylX3 is a reasonable choice from a practical standpoint, where the highest percent conversion to SpnN product is desired. The RfbB catalyzed TDP-6-deoxy-4-keto-glucose formation has been demonstrated to be a quick and efficient reaction. In addition, an excessive amount of RfbB was added in the reaction mixture to ensure that the TDP-glucose-4,6-dehydration reaction would not be rate-limiting. As shown in Figure 2-9, there were some major changes in the NMR spectrum of the TylX3/SpnN product compared to that of the starting material TDP-glucose: a sharp doublet appeared

at 1.05 ppm, there were two groups of multiplets centered at 1.70 ppm and 2.00 ppm respectively, and the portion of the spectrum between 3.20 ppm and 4.00 ppm became simplified. The spectrum is consistent with the structure of TDP-2,6-dideoxy-4-keto-glucose. The sharp doublet at 1.05 ppm was assigned as the 5-Me group; the two groups of multiplets at 1.70 ppm and 2.00 ppm were assigned to 2-H_{ax} and 2-H_{eq} respectively; the simplification of the spectrum between 3.20 and 4.00 pm was due to the disappearance of 4-H and the upfield shift of two 2-H protons. In the TDP-L-oleandrose biosynthetic pathway (Scheme 2-3), the Gra Orf27/Gra Orf26 enzyme pair catalyzes the conversion from TDP-6-deoxy-4-keto-glucose (**3**) to TDP-2,6-dideoxy-4-keto-D-glucose (**24**) whose structure has been characterized with 1D and 2D ^{1}H NMR by Floss group[70]. Our laboratory has isolated the product from the TylX3/TylC1 catalyzed reaction in the TDP-L-mycarose biosynthetic pathway (Scheme 2-3), which was established to be TDP-2,6-dideoxy-4-keto-D-allose (**28**)[71]. Although the two compounds that had been isolated in the two groups were both C-2 deoxygenated products, they were different in the stereochemistry at C-3. The C-3 hydroxyl group in the latter case was axial, as revealed by the coupling constants for 3-H and the lack of the Nuclear Overhauser Effect (NOE) between 3-H and 5-H.[71,129] In contrast, the C-3 hydroxyl group in the Gra Orf27/Gra Orf26 product is equatorial.[70] The product from the TylX3/SpnN reaction has an identical ^{1}H NMR spectrum as the compound isolated from the Gra Orf27/Gra Orf26 reaction. Combined with the evidence from the high-resolution FAB-MS, ^{31}P NMR and ^{13}C NMR, the structure of the product from the TylX3/SpnN reaction was determined to be TDP-2,6-dideoxy-4-keto-glucose (**24**) with an equatorial C-3 hydroxyl group. In view

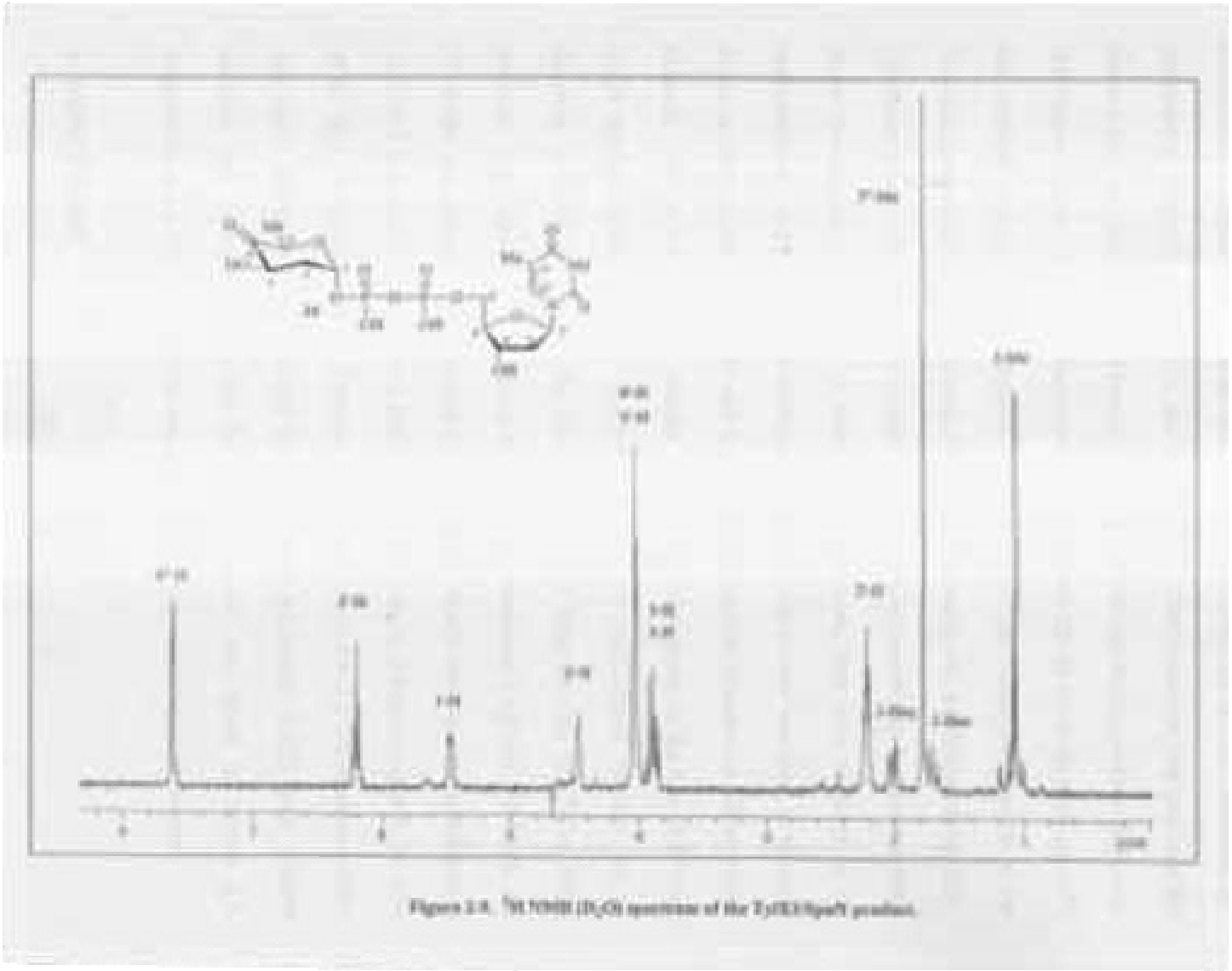


Figure 2.8. ^{1}H NMR (D_2O) spectrum of the 3'-(ES3-Spe)-T product.

of the high sequence similarity between TylX3 and SpnO, the 2,3-dehydratase activity that SpnO displayed in the activity assays, and the successful TylX3/SpnN product isolation and characterization, the functions of SpnO and SpnN as hexose 2,3-dehydratase and hexose 3-ketoreductase, respectively, were unambiguously proved.

4. CONCLUSIONS

The results presented in this chapter provide unequivocal evidence that SpnO and SpnN are TDP-6-deoxy-4-keto-D-glucose 2,3-dehydratase and TDP-2,6-dideoxy-3,4-diketo D-glucose 3-ketoreductase, respectively, in the forosamine biosynthetic pathway. Through the TylX3/SpnN product isolation and characterization, the removal of the 2-hydroxyl group in the TDP-forosamine biosynthesis is confirmed to occur through a dehydration/reduction route, providing another example of the 2-deoxygenation mechanism that has been implicated in all 2-deoxysugar biosynthetic pathways characterized thus far. This may be the general mechanism used in all sugar 2-deoxygenation reactions that occur in nature.

Chapter 3. Characterization of SpnQ, the TDP-2,6-dideoxy-4-keto-D-glucose-3-dehydrase

1. INTRODUCTION

Forosamine is a highly deoxygenated hexose which is deoxygenated at C-2, C-3, C-4 and C-6, and possesses an *N,N*-dimethylamino group at C-4. The absence of any of the hydroxyl groups on the hexose backbone distinguishes forosamine from other pyranose sugars and deoxysugars and may confer its specific biological activity. The presence of forosamine has been proven to be essential for the effectiveness of spinosad as an insectcide.[147]

The genes responsible for spinosad biosynthesis have been located and sequenced. The forosamine biosynthetic genes are clustered in a region adjacent to the polyketide synthase genes. Among them, the *spnQ*-encoded enzyme has been tentatively assigned as responsible for C-3 deoxygenation in forosamine biosynthesis based on sequence analysis. The predicted sequence of the protein is similar to those of several other sugar biosynthetic enzymes, among which CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase (E_1) has been characterized in detail. Other examples of close homologs of *spnQ* include *aknP*, in the aclarubicin biosynthesis gene cluster of *Streptomyces galilaeus*,[115] *rdmI*, in the β-rhodomycin biosynthesis gene cluster of *Streptomyces violaceus*,[177] *nanG3*, in the nanchangmycin biosynthesis gene cluster of *Streptomyces nanchangensis* NS3226,[114] and others (Table 1). Although the structures of aclarubicin, β-rhodomycin, and nanchangmycin differ significantly from each other, the

enzyme	metabolite	organism	sequence identity
RdmI	β-rhodomycin	*Streptomyces violaceus*	68.4
AknP	aclarubicin	*Streptomyces galilaeus*	67.5
NanG3	nanchangmycin	*Streptomyces nanchangensis*	64.6
UrdQ	urdamycin	*Streptomyces cyanogenus*	63.5
LanQ	landomycin	*Streptomyces fradiae*	62.6
RfbH	lipopolysaccharides	*Salmonella typhimurium*	50.2
E1	lipopolysaccharides	*Yersinia pseudotuberculosis*	48.8
ColD	lipopolysaccharides	*Yersinia pseudotuberculosis*	25.9

Table 3-1. Enzymes homologous to SpnQ . Listed are the name of the enzyme, the primary or secondary metabolite whose biosynthesis requires the enzyme, the organism where the metabolite is produced and the sequence identity of the enzyme to SpnQ.

spnQ homologs in their biosynthetic gene clusters are all proposed to encode the enzymes that catalyze C-3 deoxygenation in the biosynthesis of L-rhodinose, a 2,3-dideoxy-hexose sugar that is attached to the polyketide aglycones of these antibiotics. The C-2 deoxygenation reactions have been characterized in the biosynthesis of TDP-L-mycarose and TDP-L-oleandrose,[70,71] as have C-3 deoxygenation reactions in the biosynthesis of CDP-L-ascarylose and GDP-L-colitose.[59,60,68,69] However, none of the biosynthetic pathways of 2,3-dideoxy-hexoses including L-rhodinose have been studied in detail. As a result, the question of the order in which C-2 and C-3 deoxygenation occurs in the biosynthesis of 2,3-dideoxy-hexose remains to be answered. The two possible routes for C-2 and C-3 deoxygenation in the biosynthesis of TDP-forosamine are illustrated in Scheme 2-2. In route A, C-2 deoxygenation occurs before C-3 deoxygenation. As has

been characterized in Chapter 2, SpnO and SpnN together catalyze the removal of the C-2 hydroxyl group, producing a deoxysugar having an equatorial C-3 hydroxyl group. In the route A, this hydroxyl group is removed by the action of SpnQ. In route B, TDP-6-deoxy-4-keto-glucose is first deoxygenated at C-3 by the action of SpnQ. The reaction of SpnQ in route B is quite similar to the well-studied E_1-catalyzed reaction, with the only difference being that the NDP group attached at C-1 of the substrate is TDP rather than CDP. In route B, SpnO would catalyze hexose-2,3-dehydration after the C-3 deoxygenation to form a carbon-carbon double bond between C-2 and C-3, which would be reduced by the catalysis of SpnN. In Chapter 2, SpnN was demonstrated to be a reductase capable of catalyzing 3-ketoreduction. However, it is also possible that the 3-keto is in equilibrium with its enol form, and the reduction is carried out on the double bond. Both ketoreduction and alkene reduction are consistent with the use of NADPH as the reducing agent.

In addition to the question of reaction sequence, the C-3 deoxygenation itself is mechanistically intriguing. As described in Chapter 1, there have been two different C-3 deoxygenation mechanisms characterized so far. Although the SpnQ-catalyzed C-3 deoxygenation is not likely to follow an oxidation/reduction or dehydration/reduction way as in the characterized C-2 and C-6 deoxygenation respectively, it is possible that the reaction might have a mechanism similar to the one proposed for C-4 deoxygenation.

ColD has recently been characterized as another enzyme that is able to catalyze C-3 deoxygenation of a hexose. Like E_1, the enzyme contains a coenzyme B_6 cofactor PMP that forms a complex with the substrate at the beginning of catalysis. The ColD

reaction is different from the E_1-catalyzed reaction in that completion of the reaction does not require a reductase, but instead requires the presence of a stoichiometric amount of L-glutamate. Although SpnQ is more distantly related to ColD than to E_1, the 25.9% sequence identity between SpnQ and ColD is still significant enough as not to exclude the possibility that SpnQ catalyzes C-3 deoxygenation in a mechanism analogous to ColD. Thus mechanism A was proposed (Figure 3-1). In this mechanism, the cofactor PMP of

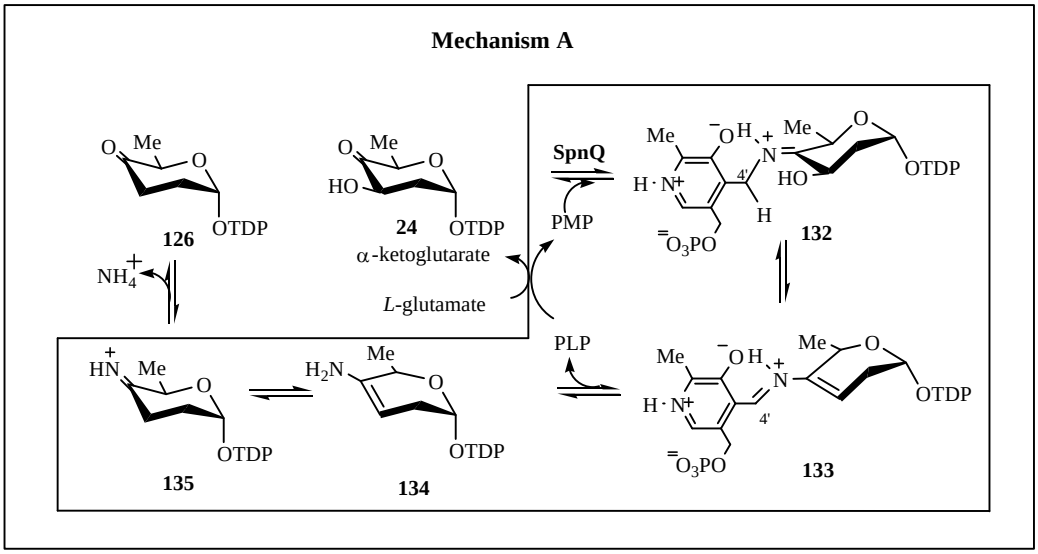


Figure 3-1 . Possible mechanism of SpnQ catalyzed C-3 deoxygenation: Mechanism A.

SpnQ forms a Schiff base with the 4-keto group of the substrate (**132**). Abstraction of a C-4' proton induces the departure of the 3-hydroxyl group of the substrate and forms a $\Delta^{3,4}$-amino-glucoseen intermediate (**133**), which upon hydrolysis generates PLP and an enamine sugar (**134**). The former reacts with L-glutamate to regenerate PMP, while the latter undergoes tautomerization followed by hydrolysis to give the 3-deoxygenated

product. The overall reaction begins with the sugar substrate and L-glutamate and ends with α-ketoglutarate, ammonia and the deoxygenated sugar product (**126**).

Another possibility is that the SpnQ-catalyzed C-3 deoxygenation has a mechanism similar to that of DesI and DesII catalyzed C-4 deoxygenation. It has been suggested that both DesI and DesII participate in the C-4 deoxygenation in the biosynthesis of desosamine either in a one-step manner requiring formation of a DesI/DesII complex or in a two-step manner. DesI is proposed to be responsible for the aminotransfer reaction that occurs at C-4, while DesII is suggested to be a SAM radical dependent iron-sulfur protein catalyzing deamination at C-4 and oxidation at C-3. Most of SAM radical-dependent proteins have a consensus CXXXCXXC motif that is proposed to be the [4Fe-4S] binding site. SAM is suggested to bind near the iron-sulfur center. Induced by an electron transferred from the iron-sulfur center, a molecule of SAM undergoes reductive cleavage to produce a molecule of methionine and an adenosyl radical that is capable of initiating complicated reactions. From sequence analysis, SpnQ is homologous to quite a number of aminotransferases and has the four conserved cysteine residues that are the anchors of the iron-sulfur center in E_1. The iron-sulfur center in E_1 has been well characterized as a [2Fe-2S] center.[55,60,64,65] Based on the overall sequence similarity and the perfect alignment of conserved cysteines between SpnQ and E_1, SpnQ has been proposed to contain an [2Fe-2S] as well. However, a [4Fe-4S] is not an absolute requirement for a protein to be included in the SAM radical-dependent superfamily.[132] Thus, SpnQ is a possible SAM-radical dependent protein candidate. Taken together, this information indicates that SpnQ could act similarly to the

combination of DesI and DesII, and as a result could catalyze C-3 deoxygenation following a mechanism similar to that of DesI and DesII catalyzed C-4 deoxygenation. In order to get a C-3 deoxygenated product by the mechanism just described, the substrate has to first undergo isomerization to convert TDP-2,6-dideoxy-4-keto-glucose (**24**) to TDP-2,6-dideoxy-3-keto-glucose (**136**). An enzyme catalyzing such an isomerization has been identified in the biosynthetic gene cluster of TDP-mycaminose.[153] However, this isomerization can also occur as a nonenzymatic process.[30,31,129] This could explain why an isomerase gene has not been discovered in the biosynthetic gene cluster of TDP-forosamine. Thus, another mechanism, mechanism B (Figure 3-2), of SpnQ-catalyzed C-3 deoxygenation is proposed. In mechanism B, the substrate, TDP-2,6-dideoxy-4-keto-glucose (**24**)

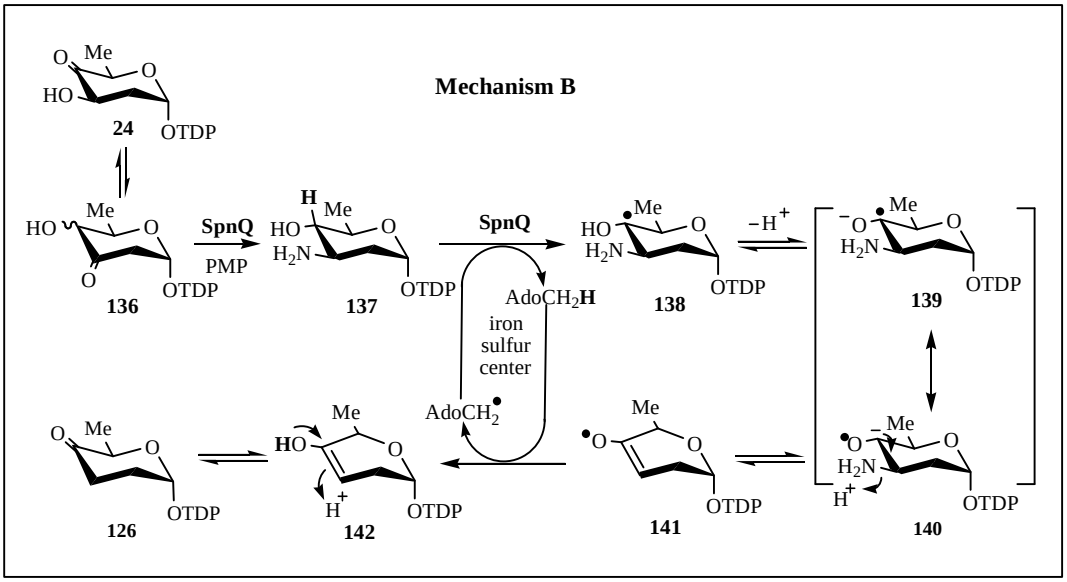


Figure 3-2 . Possible mechanism of SpnQ catalyzed C-3 deoxygenation: Mechanism B.

is first converted to TDP-2,6-dideoxy-3-keto-glucose (**136**). Then SpnQ catalyzes the formation of TDP-2,6-dideoxy-3-amino-glucose (**137**). The SAM molecule that is proposed to bind to SpnQ undergoes reductive cleavage to produce an adenosyl radical. The active adenosyl radical then abstracts a hydrogen from C-4 of the sugar to form a molecule of adenosine and a radical at C-4 (**138**). After deprotonation, C-4 becomes a carbanion while the radical is relocated to the C-4 oxygen (**140**). The high energy C-4 carbanion drives the elimination of the amino group at C-3 (**141**). Next, the oxygen radical abstracts a hydrogen from the adenosine to yield an enol-hexose (**142**), which upon tautomerization transforms to TDP-2,3,6-trideoxy-4-keto-glucose (**126**) and regenerating the adenosyl radical for the next catalysis cycle. It should be noted that since the SAM radical-catalyzed deoxygenation is still mechanistically unsolved, there are several ways to illustrate the involvement of the SAM radical. Mechanism B is only one of them.

In the E_1-catalyzed reaction, the completion of C-3 deoxygenation is assisted by another protein, CDP-6-dexoy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase (E_1) reductase (E_3). The coenzyme B_6 cofactor PMP in E_1 reacts with the substrate to form a CDP-6-deoxy- $\Delta^{3,4}$-glucoseen intermediate (**70**, Figure 1-10) that is reduced in a step-wise manner by the electrons transferred from NADH through E_3. E_3 is a ferredoxin reductase-like protein containing a [2Fe-2S] cluster and a bound FAD cofactor. Considering the high sequence identity between E_1 and SpnQ (49% identity), SpnQ was originally assumed to catalyze the C-3 deoxygenation in a manner analogous to E_1.

Accordingly, mechanism C was proposed (Figure 3-3). However, an ORF that encodes an E_3-like protein has not been found in the biosynthetic gene cluster of spinosyn. As has

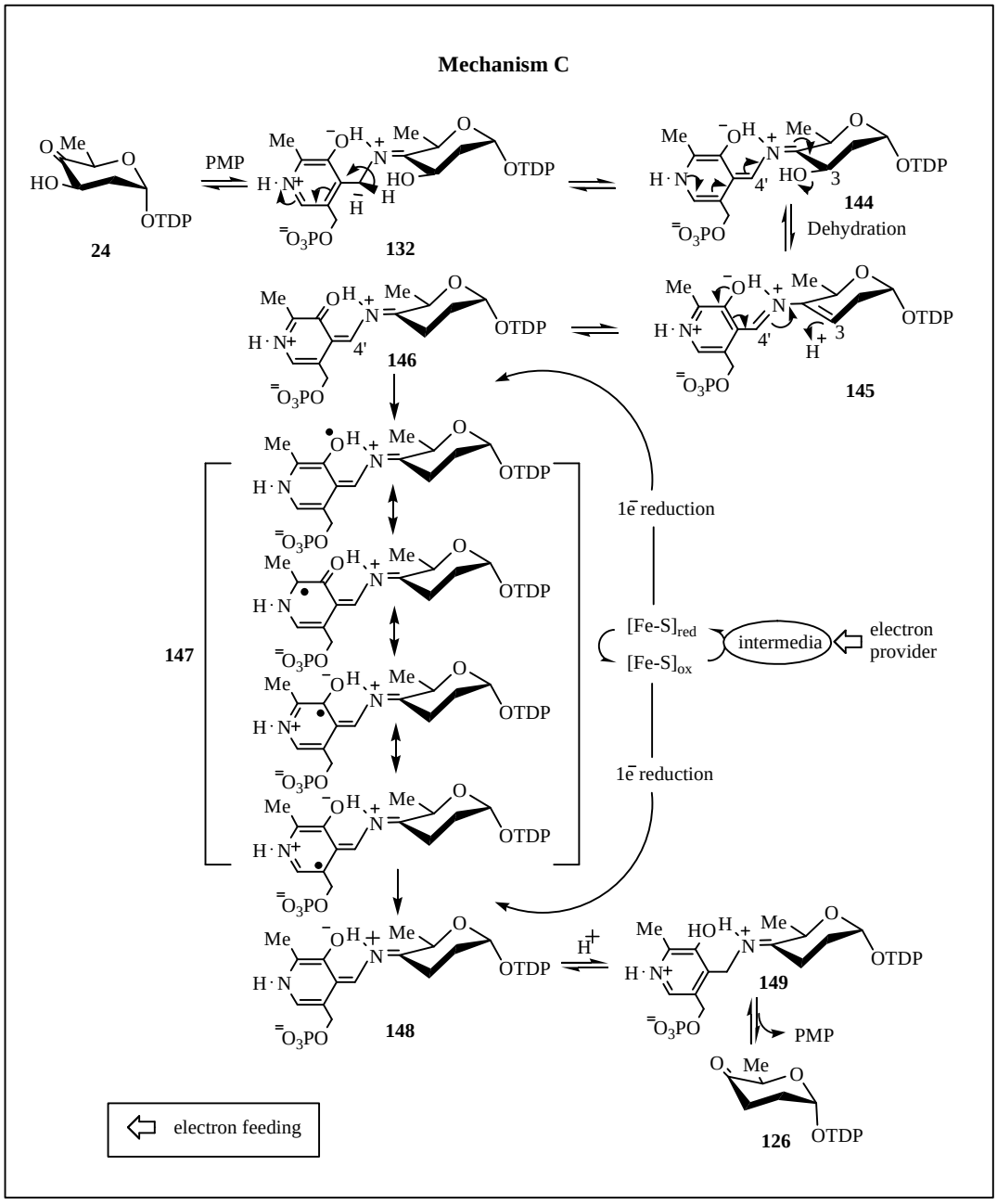


Figure 3-3 . Possible mechanism of SpnQ catalyzed C-3 deoxygenation: Mechanism C.

been illustrated in Chapter 2, the genes that are involved in the biosynthesis of forosamine are clustered together and the protein encoded by each of them has been assigned with an appropriate function based on sequence analysis. It is highly unlikely that a gene located in the region encoding the polyketide synthase (PKS) could beinvolved in the C-3 deoxygenation since the genes for type-I PKS are present as very large ORFs and the domains within each ORF have all been assigned functions. It is also highly unlikely that the genes in the biosynthetic gene cluster that have not been assigned with any functions might encode the auxiliary protein of SpnQ. Sequence analysis shows that none of them shares significant sequence identity with E_3. Moreover, the biosynthetic gene clusters of all the previously mentioned antibiotics whose structures contain an attached L-rhodinose have also been scrutinized. None of them contains an ORF whose encoded protein has significant sequence similarity to E_3. Thus it is puzzling how SpnQ might catalyze the reaction in a manner analogous to E_1/E_3, but without a reductase.

In this chapter, the expression and purification of SpnQ are described. SpnQ was shown *in vitro* to be the enzyme catalyzing C-3 deoxygenation of the SpnO/SpnN product. The reaction requires the electron transfer protein pair, either ferredoxin/ferredoxin reductase or flavodoxin/flavodoxin reductase. The mechanism of the SpnQ catalyzed C-3 deoxygenation is likely to be similar to the E_1/E_3 mechanism, with the combination of ferredoxin or flavodoxin and its respective reductase playing the role of E_3 and SpnQ acting analogously to E_1. Since no reaction occurred when SpnQ was incubated with TDP-6-deoxy-4-keto-glucose (**3**) under identical condition used to

show SpnQ activity with the SpnN product (**24**), it was shown that C-2 deoxygenation precedes C-3 deoxygenation.

2. EXPERIMENTAL PROCEDURES

Materials. Most materials used in the experiments in this chapter were the same as those described in Chapter 2. Ferredoxin and Ferredoxin reductase were purchased from Sigma-Aldrich. The E_1 protein was kindly provided by Dr. Beth Paschal in our laboratory. *E. coli JM105*-pOPI frozen culture, used for overexpression and purification of E_3, was obtained from central collection in our laboratory. Econo-Pac 10 DG column was purchased from Bio-Rad. Any chemicals that were used in this chapter but not mentioned in Chapter 2 were purchased from Sigma.

General Equipment. The equipment and other experimental apparatus were the same as those described in Chapter 2.

Bacterial Strains. Various *E. coli* host cell strains used were the same as those described in Chapter 2.

PCR Amplification of the spnQ Gene. Three oligonucleotide primers were obtained to amplify the *spnQ* gene: primer (Q-1) 5'-GGGGCCTC**CATATG**AGCAGT-TCTGTCGAAGC-3' containing an *Nde*I restriction site (in bold), primer (Q-2) 5'-GAG**CTCGAG**TTATCGCCCCAACGCCCACAAG-3' containing an *Xho*I restriction site (in bold), and primer (Q-3) 5'-TAT**CTCGAG**TCGCCCCAACGCCCACAAGC-3' containing an *Xho*I restriction site (in bold). The start primer Q-1 was designed to hybridize to the first 20 nucleotides at the 5' end of *spnQ* gene while the halt primers Q-2

and Q-3 are designed to hybridized to 23 and 20 nucleotides at the 3' end of the gene, respectively. Two sets of polymerase-mediated amplifications were carried out, both in 0.5 mL microcentrifuge tubes. In the first reaction, 37.5 μL of double-deionized H_2O, 5.0 μL of *pfu* polymerase buffer (10 x), 4.0 μL of deoxyribonucleotidyltriphosphate mix (2.5 mM each), 1.0 μL of start primer Q-1 (10 μM), 1.0 μL of halt primer Q-2 (10 μM), 0.5 μL of genomic DNA (about 0.1 μg) from *Saccharopolyspora spinosa* as the amplification template, and 1.0 μL of cloned *pfu* polymerase (2.5 units) were used. The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following thermal cycles: (1) 1 denaturation cycle of incubation at 95 °C for 3 min; (2) 30 amplification cycles of incubation at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 2.5 min; (3) 1 extension cycle of incubation at 72 °C for 10 min. The reaction tube was held at 4 °C prior to further manipulations. This amplification product of the *spnQ* gene containing its natural stop codon, was called Q-w-stop. In the second polymerase mediated amplification reaction, all components were the same as in the first reaction except that the halt primer was changed to Q-3 of same concentration. The thermal cycles and the treatment of the second reaction after thermal cycling were the same as in the first reaction. The stop codon of the natural *spnQ* gene was deleted in the product of the second amplification reaction. This PCR product was named Q-wo-stop.

Cloning of SpnQ Gene. After PCR amplification, standard recombinant DNA techniques were used to clone the genes into pET24b(+). Ligation of Q-w-stop into pET24b(+) gave pKZQO, and ligation of Q-wo-stop into pET24b(+) gave pKZQC. PKZQO, when expressed, should produce the native SpnQ protein, while pKZQC, when

expressed, should produce the SpnQ protein with a *C*-terminal tail of about 8 amino acids containing a His_6 repeat at the extreme *C*-terminal end. Positive clones were identified by digestion analysis of the plasmid DNA with the appropriate restriction enzymes and further confirmed by sequencing the plasmid DNA with T7 promoter and T7 terminator primers (Novagen). The plasmid DNAs of the positive clones, pKZQC and pKZQO were used to transform *E. coli* BL21(DE3). The general methods and protocols for recombinant DNA manipulations used were as described by Sambrook *et al*.[156]

Purification of C-terminal His-tagged SpnQ. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli-pKZQC Cells. An overnight culture of *E. coli* BL21(DE3)-pKZQC, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was then adjusted to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 18 h at 24 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with one minute cooling intervals between bursts. Cellular debris was removed by centrifugation

at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract was added 10 mL of Ni-NTA slurry (5 mL of resin) and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The resin was then spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g) and the majority of the supernatant was decanted. The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 25 mL of lysis buffer and various amounts of wash buffers. All the wash buffers and the elution buffer used contained the same ingredients as the lysis buffer except that each contained a different concentration of imidazole. The column was washed sequentially with 50 mL of wash buffer containing 20 mM imidazole, 20 mL of wash buffer containing 40 mM imidazole, and 20 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 0.8 mL/min. After washing, the SpnQ protein was eluted with 10 mL of the elution buffer containing 250 mM imidazole. Eluted protein was transferred into a dialysis bag and dialyzed against four 1 L portions of dialysis buffer (50 mM Tris·HCl, 15% glycerol, pH 7.5) for 1 h per liter. The dialyzed protein was concentrated by ultrafiltration using an Amicon concentrator and a YM-10 membrane (Amicon) and stored at –80 °C.

Purification of Native SpnQ. All purification steps were carried out at 4 °C.

Step 1: Growth of E. coli BL21 (DE3)-pKZQO Cells. An overnight culture of *E. coli* BL21(DE3)-pKZQO, grown in LB medium supplemented with kanamycin

(30μg/mL) at 37 °C, was diluted 250-fold with 6 L of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was adjusted to 24 °C and the incubation was continued for one more hour in order for the OD_{600} of the culture to reach 0.4. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 18 h at 24 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. Thawed cells were resuspended in 60 mL of lysis buffer (25 mM Tris·HCl, 0.6 mM PMSF, 5 mM 2-mercapto-ethanol, 10% glycerol, pH 7.5), and disrupted with five 1 min sonication bursts, with 1 min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 25 min.

Step 3: Ammonium Sulfate Precipitation. Ammonium sulfate crystals were added to the supernatant from the previous step in two portions over 20 min to give 10% saturation. After stirring for 1 h, the precipitated proteins were removed by centrifugation at 15,000 g for 20 min and more ammonium sulfate crystals were added to the supernatant in three portions over 30 min to give 70% saturation. The mixture was slowly stirred for 2 h. The precipitated protein was collected by centrifugation at 15,000g for 20min, and resuspended in a minimal amount of lysis buffer. The resulting protein solution was desalted by dialysis against four 1 L of lysis buffer which were changed every 1.5 h.

Step 4: DEAE-Sepharose CL6B Chromatography. The dialysate from step 3 was applied to a DEAE-Sepharose CL6B column (2.5 x 40 cm) preequalibrated with buffer A

(25 mM Tris·HCl, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The column was washed with 200 mL of the same buffer, and eluted with a linear gradient of 800 mL of buffer A and 800 mL of buffer B (buffer A plus 0.8 M NaCl). The flow rate was 1.5 mL/min and fractions of 12 mL were collected throughout the entire gradient elution. The fractions containing the desired SpnQ protein as judged by SDS-PAGE were combined, desalted by repeated dilution with buffer A and concentration by ultrafiltration, and finally concentrated to 8 mL.

Step 5: Hydroxyapatite Chromatography. The protein from the previous step was applied to a Hydroxyapatite column (2.5 x 22 cm) that was preequalibrated with buffer C (10 mM potassium phosphate, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The column was washed with 100 mL of the same buffer, and eluted with a linear gradient of 400 mL of buffer C and 400 mL of buffer D (600 mM potassium phosphate, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The flow rate was 1.0 mL/min and fractions of 8 mL were collected throughout the entire gradient elution. The fractions containing the desired SpnQ protein as judged by SDS-PAGE, were pooled and concentrated to 3 mL by ultrafiltration.

Step 6: Sephacryl S-200 HR Chromatography. The protein from step 5 was applied to a Sephacryl S-200 HR column (2.5 x 100 cm) preequalibrated with buffer E (25 mM Tris·HCl, 15% glycerol, pH 7.5). The column was washed/eluted with the same buffer at a flow rate of 0.8 mL/min. Fractions of 8 mL were collected throughout the entire elution. Pure fractions were identified by SDS-PAGE, and concentrated to 2.0 mL by ultrafiltration. The purified SpnQ protein was stored at –80 °C.

Purification of E_3. All purification steps were carried out at 4 °C, except for the FPLC step, which was done at room temperature.

Step 1: Growth of E. coli JM105-pOPI *Cells.* An overnight culture of *E. coli* JM105-pOPI, grown in LB medium supplemented with ampicillin (100μg/mL) at 37 °C, was diluted 250-fold with 6 L of TB (terrific broth) medium containing 100 μg/mL ampicillin and incubated at 37 °C until the OD_{600} reached 0.6. The culture was then induced with IPTG to a final concentration of 0.25 mM and allowed to grow for an additional 12 h at 37 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 12 g per liter.

Step 2: Crude Extract Preparation. Thawed cells were resuspended in 60 mL of lysis buffer (25 mM Tris·HCl, 0.6 mM PMSF, 5 mM 2-mercapto-ethanol, 0.1 mM EDTA, 10% glycerol, pH 7.5), and disrupted with six 1 min sonication bursts, with 1 min cooling intervals between bursts. Cell debris was removed by centrifugation at 35,000 g for 25 min.

Step 3: Ammonium Sulfate Precipitation. Ammonium sulfate crystals were added in three portions over 30 min to the protein solution from the previous step to give 25% saturation. After stirring for 0.5 h, the precipitated proteins were removed by centrifugation at 15,000 g for 20 min and more ammonium sulfate crystals were added in three portions over 30 min to give 72% saturation. The mixture was slowly stirred for 0.5 h. The precipitated protein was collected by centrifugation at 15,000g for 20 min, and resuspended in a minimal amount of lysis buffer. The resulting protein solution was desalted by dialysis against four 1 L of lysis buffer over the course of 5 h.

Step 4: DEAE-Sepharose CL6B Chromatography. The dialysate from step 3 was applied to a DEAE-Sepharose CL6B column (2.5 x 40 cm) preequalibrated with buffer A (25 mM Tris·HCl, 1 mM 2-mercapto-ethanol, 0.1 mM EDTA, 15% glycerol, pH 7.5). The column was washed with 200 mL of the same buffer, and eluted with a linear gradient of 600 mL of buffer A and 600 mL of buffer B (buffer A plus 0.3 M NaCl). The flow rate was 1.5 mL/min and fractions of 12 mL were collected throughout the entire gradient elution. The fractions containing the desired E_3 protein as judged by SDS-PAGE were combined and ammonium sulfate was added to give 72% saturation. The ammonium sulfate precipitated proteins were collected by centrifugation at 15,000g for 20 min and resuspended in a minimal amount of buffer C (buffer A plus 3.0 M NaCl).

Step 5: Phenyl-Sepharose Chromatography. The protein solution from the previous step was applied onto a phenyl-Sepharose column (2.5 x 18 cm) preequalibrated with buffer C. The column was washed with 100 mL of the same buffer and eluted with a linear gradient of 250 mL of buffer C and 250 mL of buffer A. The flow rate was 1.5 mL/min and fractions of 12 mL were collected throughout the entire gradient elution. The fractions containing the desired E_3 protein as judged by SDS-PAGE, were pooled, and desalted by repeated dilution with buffer A and concentration by ultrafiltration, and finally concentrated to 3 mL.

Step 6: FPLC MonoQ Chromatography. The protein from step 5 was further purified by FPLC using a MonoQ HR (10/10) column with buffer A and buffer D (buffer A plus 0.5 M NaCl). The elution profile began with a linear gradient of 0 to 100% B in 20 min followed by a wash with 100% B for 6 min and a reequalibration with buffer A

for 6 min. The flow rate was 3 mL /min and the detector was set to 280 nm. The sharp peak with a retention time of 10.0 min was collected, desalted by repeated dilution with storage buffer (25 mM Tris·HCl, 15% glycerol, pH 7.5) and concentration by ultrafiltration, and finally concentrated to 3 mL and stored at –80 °C. A typical yield of 120 mg of pure E_3 was obtained from 6 L culture.

Purification of FLD Protein. The gene encoding flavodoxin (FLD) has been cloned into vector pET24b(+)by Dr. Hua Zhang. The resulting expressed protein will have a *C*-terminal tail containing a His_6 tag at the extreme *C*-terminus. All purification operations were carried out at 4 °C.

Step 1: Growth E. coli BL21-pHZFLD cells. An overnight culture of *E. coli* BL21(DE3)-pHZFLD, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold with 2 L of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The culture was then induced with IPTG to a final concentration of 0.1 mM and allowed to grow for an additional 6 h at 37 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 8 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 50 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with three 1 min sonication bursts, with 1 min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. A volume of 6 mL of Ni-NTA slurry (3 mL of resin) was added to the crude extract and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 80 mL of wash buffer and eluted with 15 mL of elution buffer. The wash buffer and elution buffer have the same components as the lysis buffer except that they contain 20 mM and 250 mM imidazole, respectively. The buffer passed through the column by gravity at a flow rate of 1.0 mL/min. The eluted flavodoxin protein was transferred into a dialysis bag and dialyzed against four 1 L portions of dialysis buffer (50 mM Tris, 15% glycerol, pH 7.5), changing the buffer every hour. The dialyzed protein was concentrated by ultrafiltration using an Amicon concentrator with an YM-10 membrane (Amicon) and stored at –80 °C.

Purification of FNR Protein. The gene encoding flavodoxin reductase (FNR) has been cloned into vector pET24b(+) by Dr. Hua Zhang in our laboratory. The resulting expressed protein has a *C*-terminal tail containing a His_6 tag at the extreme *C*-terminus. The purification procedure of FNR is same as that of FLD. The frozen culture used to initiate cell culture was *E. coli* BL21-pHZFNR cells from central collection of our laboratory.

Apo-enzyme Preparation. The as-purified native SpnQ protein was mixed with $Na_2S_2O_4$ and EDTA in a molar ratio of 1.0: 12.2: 16.2 in 1.5 mL of Tris·HCl buffer (25 mM, 15% glycerol, pH 7.5). The resulting solution was incubated at room temperature

(24 °C) for 1.5 h. The excess $Na_2S_2O_4$ and EDTA were removed by Econo-Pac 10 DG column (Bio-Rad) according to the provided protocols. Briefly, the reaction mixture was diluted to 3 mL with Tris·HCl buffer and was applied to the column which had been equilibrated with Tris·HCl buffer (25 mM Tris·HCl, 15 % glycerol, pH 7.5). The column was washed with the same buffer and the first 3 mL of eluent from the column was collected as the desalted protein. The protein was concentrated by ultrafiltration and stored at –80 °C.

Reconstitution of SpnQ protein. Reconstitution of native SpnQ protein or apo-SpnQ protein was conducted in an anaerobic environment. All the buffer, chemicals and enzyme were degassed according to standard procedures before being put into the anaerobic glove box (Coy Laboratory Products Inc.). The reconstitution was performed in the glove box under the atmosphere which has an O_2 content lower than 5 ppm. DTT was added to the enzyme solution at a final concentration of 10 mM and the resulting solution was incubated for 20 min. Then, 5 molar equivalents of PMP, PLP, $Fe(NH_4)_2(SO_4)_2$, and Na_2S, compared to the amount of SpnQ, were sequentially added to the enzyme solution with a 15 min incubation period between each addition. The addition of $Fe(NH_4)_2(SO_4)_2$ and Na_2S, and the 15 minute incubation intervals, were repeated until the final concentrations of $Fe(NH_4)_2(SO_4)_2$ and Na_2S each reached 20 times that of SpnQ. The extra salts were removed using an Econo-Pac 10 DG column as described previously. The reconstituted protein was concentrated with Centricon YM-10 by microcentrifugation.

Polyacrylamide Gel Electrophoresis. The subunit molecular mass and the purity of the protein samples were assessed by SDS-PAGE as described in Chapter 2.

Molecular Weight Determination. The molecular weight of the native SpnQ was determined by gel filtration chromatography as described in Chapter 2.

N-Terminal Amino Acid Analysis. The purified native SpnQ protein were submitted to the Institute for Cellular and Molecular Biology at UT, Austin for *N*-terminal amino acid analysis.

UV-visible Scan of SpnQ Protein. The absorbance of His_6-tagged SpnQ, native SpnQ, and reconstituted SpnQ were scanned over the range of 250 nm to 750 nm on a Beckman DU-650 spectrophotometer.

Iron Quantitation of SpnQ Protein. Iron determination was accomplished by a well established method.[329] A 1 mL solution containing either as-isolated or reconstituted SpnQ (1 mg) was treated with 0.5 mL of freshly prepared reagent A (2.25% w/v $KMnO_4$ in 6 N HCl) at 60 °C for 2 h. After the treatment was complete, 100 μL of freshly prepared reagent B (6.5 mM ferrozine, 13.1 mM neocuprine, 2 M ascorbic acid, and 5 M ammonium acetate in 25 mL ddH_2O) was added, followed by vigorous mixing and incubation for 30 min at room temperature. A number of standards made by different concentrations of $Fe(NH_4)_2(SO_4)_2$ solution in 0.01 M HCl, were treated in parallel with the protein sample, and the iron content of SpnQ was determined by comparing the absorbance of the SpnQ sample at A_{562} to the standard curve.

The enzymatic activity of SpnQ was examined under three conditions: condition A, condition B, and condition C, were used to test whether the SpnQ-catalyzed C-3

deoxygenation follows mechanism A, mechanism B, or mechanism C, respectively. Briefly, mechanism A is a ColD-like mechanism, and condition A contained both PLP and L-glutamate and was carried out under an aerobic atmosphere. Mechanism B is a DesI/DesII-like mechanism, and condition B was conducted in the presence of dithionite under an anaerobic atmosphere. Mechanism C is an E_1/E_3-like mechanism, and condition C included an electron transfer intermediary and a reducing agent carried out in an aerobic atmosphere.

Activity Assay for SpnQ - Condition A. The assay was developed for assessing SpnQ activity by monitoring the disappearance of the SpnN product (**24**) and the concomitant formation of new peak(s) at 267 nm with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) equipped with a CarboPac™ PA1 guard column (4 x 50 mm). A typical assay mixture of 100 μL contained 1.05 mM of the SpnN product, 0.25 mM PLP, 3.0 mM L-glutamate, appropriate amounts of SpnQ, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of the enzyme and was carried out at 37 °C. A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by 10 fold dilution of the sample with water and flash freezing in liquid nitrogen. After being redissolved at 4 °C, the enzyme was removed from the solution by filtering through Centricon YM-10 membrane by centrifugation and an appropriate amount of the filtrate was subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60% B over 20 min between Buffer A (water) and Buffer B (500 mM NH_4OAc, pH 7.0 adjusted with diluted $NH_3 \cdot H_2O$) gave baseline separation of the SpnN product

(retention time = 33.4 min) and a new SpnQ reaction product (retention time = 14.7 min). The elution gradient was followed by a final wash with 100% B for 5 min and a reequilibration with 5% B for 15 min. The absolute rate was calculated as the number of nmols of product formation per min and percent conversion of the product was used directly for the relative rate.

Preparation and Characterization of the SpnQ Product - Condition A. A 200 μL incubation contained 0.21 μmol of the SpnN product, 0.05 μmol of PLP, 0.6 μmol of L-glutamate, 0.01 μmol of SpnQ in 50 mM potassium phosphate buffer (pH 7.5). The reaction was conducted at 37 °C overnight. The enzyme was removed by filtering through Centricon YM-10 by microcentrifugation. Aliquots of 20 μL of the filtrate were repeatedly injected into HPLC and separated with the Dionex CarboPac^{TM} PA1 analytical column. The reaction product whose retention time was 14.7 min was manually collected. After all the reaction filtrate was injected, the collected eluent from each individual run was combined and lyophilized. The sample was then redissolved in 100 μL of H_2O and subjected to ESI-MS analysis.

Determination of the Amino Donor for the SpnQ reaction - Condition A. The efficiency of two amino acids, L-aspartate and L-glutamate as amino donors in the SpnQ catalyzed reaction were examined. The amino acids were added to a final concentration of 3 mM in an SpnQ activity assay as previously described. The relative efficiencies of these two amino acids as amino donors in the SpnQ reaction were determined by comparing the relative reaction rate with each co-substrate.

Determination of the Optimal Temperature for the SpnQ Reaction - Condition A. The optimal temperature for the reaction catalyzed by SpnQ was investigated using the previously described assay, but varying the temperature. Temperatures of 37 °C, 29 °C, or 24 °C were used. The relative efficiencies of the reaction at each temperature were determined by comparing the relative reaction rate at each temperature.

Determination of the Optimal pH for the SpnQ Reaction – Condition A. The optimal pH for the reaction catalyzed by SpnQ was examined using the previously described assay except that different buffers were used to obtain different pH values in the reaction mixture: a 50 mM MES buffer was used for the pH 5.5 reaction; a 50 mM Bis-Tris-propane buffer was used for the pH 6.5 reaction; a 50 mM potassium phosphate buffer was used for the pH 7.5 reaction; and a 50 mM Bis-Tris-propane buffer was used for the pH 8.5 and pH 9.5 reactions. The relative reaction rates at each pH value were compared to determine the pH optimum.

The Effect of SAM on the SpnQ Reaction –Condition A. To examine the possibility that SAM might play a role in the SpnQ-catalyzed reaction, SAM was included in an assay using the previously described conditions at a final concentration of 200 μM. The relative rates of the reactions in the presence versus absence of SAM were compared.

The Effect of DTT on the SpnQ Reaction – Condition A. DTT was included in an activity assay at a final concentration of 5 mM, and the relative reaction rate in the presence of DTT was compared with that in its absence.

The Effect of SpnQ Reconstitution on SpnQ Activity - Condition A. The same amounts of reconstituted and as-isolated native SpnQ were used in activity assays performed in parallel, and the relative rates of the two reactions were compared.

The Effect of SpnQ His-tag on SpnQ Activity -Condition A. His-tagged SpnQ, in stead of the native SpnQ, was included in the activity assays. The relative rates of the reactions using His-tagged SpnQ and native SpnQ were compared.

Activity Assay for SpnQ - Condition B. Reconstituted SpnQ obtained as described previously was treated with 20 equivalents of dithionite and was incubated at room temperature for 40 min in the anaerobic glove box in order to fully reduce the iron-sulfur clusters in preparation for the assay. The resulting protein solution was used in the activity assay described below.

An assay was developed for SpnQ by monitoring the disappearance of the SpnN product and the concomitant formation of new peak(s) at 267 nm with a Dionex CarboPac^{TM} PA1 analytical column (4 x 250 mm) connected to a CarboPac^{TM} PA1 guard column (4 x 50 mm). A typical assay mixture of 100 μL contained 0.60 mM of SpnN product, 0.25 mM PMP, 0.60 mM dithionite, 30 μM SpnQ, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of SpnQ and was carried out at room temperature (24 °C). A suitable amount of the reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by 10 fold dilution with water and flash freezing in liquid nitrogen. After being redissolved at 4 °C, the enzyme was removed by filtering through Centricon YM-10 membrane by centrifugation and an appropriate amount of the filtrate was subjected to HPLC analysis. A linear

gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60% B over 20 min between Buffer A (water) and Buffer B (500 mM NH_4OAc, pH 7.0 adjusted with diluted $NH_3 \cdot H_2O$) gave base-line separation of the SpnN product (retention time = 33.4 min) and a new SpnQ reaction product (retention time = 35.8 min). The elution gradient was followed by a final wash with 100% B for 5 min and a reequilibration with 5% B for 15 min. The absolute rate was calculated as the number of nmols of product formed per min and percent conversion of the product was used to directly determine the relative rate.

The Effect of $Na_2S_2O_4$ Concentration on SpnQ Activity - Condition B. To examine whether the SpnQ reaction can sustain a high concentration of $Na_2S_2O_4$, an assay was conducted using previously described conditions except that 2.5 mM $Na_2S_2O_4$ was included. The relative reaction rate was compared to that under standard condition B.

The Effect of SAM on SpnQ Activity - Condition B. SAM was included in the assays described previously at a concentration of 0.3 mM, which is equal to 5 molar equivalents of SpnQ used, or 1.2 mM, which is equal to 2 molar equivalents of substrate used. The relative rates of the reactions were compared to that under standard condition B without SAM to determine the effect of SAM on the catalytic efficiency of SpnQ.

Activity Assay for SpnQ – Condition C. An as-isolated SpnQ was used in this assay. The assay was developed for SpnQ by monitoring the disappearance of the SpnN product and the concomitant formation of new peak(s) at 267 nm with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) connected to a CarboPac™ PA1 guard column (4 x 50 mm). A typical assay mixture of 50 μL contained 0.70 mM of SpnN

product, 0.15 mM PMP, 30 μM ferredoxin, 30 μM ferredoxin reductase, 4.0 mM NADPH, 30 μM SpnQ, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of NADPH and was carried out at room temperature (24 °C). A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by 10 fold dilution with water and flash freezing in liquid nitrogen. After being dissolved at 4 °C, the enzyme was removed by filtering through Centricon YM-10 membrane by centrifugation and an appropriate amount of the filtrate was subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60% B over 20 min between Buffer A (water) and Buffer B (500 mM NH_4OAc, pH 7.0 adjusted with diluted $NH_3 \cdot H_2O$) gave base-line separation of the SpnN product (retention time = 33.4 min), NADP (retention time = 39.2 min), and a new SpnQ reaction product (retention time = 35.8 min). The elution gradient was followed by a final wash with 100% B for 5 min and a reequilibration with 5% B for 15 min. The absolute rate was calculated as the number of nmols of product formed per min and percent conversion of the product was used to directly determine the relative rate.

Substrate Specificity of SpnQ – Condition C. TDP-6-deoxy-4-keto-glucose (**3**) was examined as a possible substrate for SpnQ. Using the HPLC assay described previously, 1.0 mM of TDP-6-deoxy-4-keto-glucose, instead of the SpnN product (**24**), was included in 50 μL of 50 mM potassium phosphate buffer (pH 7.5), along with 1.5 mM NADPH, 0.15 mM PMP, 30 μM ferredoxin, 30 μM ferredoxin reductase, and 30 μM SpnQ. At certain time intervals, an aliquot of 5 μL was withdrawn and analyzed for new

product formation. After the reaction was allowed to proceed for 2 h at room temperature, all the enzymes were removed from the reaction mixture by filtering through Centricon YM-10 membrane by centrifugation. To the supernatant was added 1.0 mM NADPH, 30 μM SpnN and 20 μM SpnO. After another two hours, the reaction was analyzed by HPLC following the procedures described for the SpnQ activity assay.

Testing the Activity of the E_1/E_3 Pair Towards the SpnN Product. In view of the significant sequence similarity between E_1 and SpnQ, it would be interesting to examine whether the SpnQ substrate (**24**) can be converted to the C-3 deoxygenated product by the action of E_1 and E_3. This possibility was tested by an assay similar to the SpnQ activity assay under condition C. A typical assay mixture of 25 μL contained 1.05 mM of the SpnN product, 0.30 mM PMP, 60 μM E_3, 2.0 mM NADH, 30 μM E_1, and 50 mM potassium phosphate (pH 7.5). It should be noted that in this reaction mixture, NADH, instead of NADPH, was used because NADH is the well-established reducing agent in the E_1/E_3 system. The reaction was initiated by the addition of NADH and was carried out at 24 °C. At appropriate time intervals, a suitable amount of the reaction mixture was subjected to HPLC analysis after removing the enzyme by Centricon YM-10. The HPLC program used to separate the compounds was same as that used for the SpnQ activity assay (condition C).

Assay of E_3 as the Electron Transfer Mediator in the SpnQ Catalyzed Reaction–Condition C. As has been explained previously, a gene encoding an E_3 homologue was not found in the TDP-forosamine biosynthetic gene cluster. However, this could be due to the inherent limitations of using gene sequence as an identification method. To

examine whether the SpnQ catalyzed reaction needs an E_3-like reductase, E_3 was included in the assay mixture. A typical reaction of 25 μL contained 1.05 mM of the SpnN product, 0.30 mM PMP, 60 μM E_3, 2.0 mM NADH, 30 μM SpnQ, and 50 mM potassium phosphate (pH 7.5). NADH is the reducing agent in the E_1/E_3 reaction, providing electrons to E_3. A suitable amount of reaction mixture was withdrawn at appropriate time intervals, was treated in the same way as in the SpnQ activity assay, and was subjected to HPLC analysis. The same elution program as has been previously described for the SpnQ activity assay was used to separate the compounds.

Testing the Possibility that SpnQ Uses NADH, FAD as the Electron Source - Condition C. An assay mixture of 25 μL contained 1.05 mM of SpnN product, 0.30 mM PMP, 600 μM FAD, 2.0 mM NADH, 30 μM SpnQ, and 50 mM potassium phosphate (pH 7.5). It should be noted that the reducing agent used is NADH. A suitable amount of reaction mixture was withdrawn at appropriate time intervals, treated in the same way as in the SpnQ activity assay, and subjected to HPLC analysis. The same elution program that was described for previous SpnQ assays was used to separate the compounds.

Determining the Effect of Enzyme Reconstitution on the SpnQ/E_3 (FAD) Reaction. To examine the effect of reconstitution of SpnQ on the SpnQ/E_3 (FAD) reaction, the as-isolated native SpnQ was replaced with reconstituted SpnQ in the reactions. The reaction contained the SpnN product, PMP, NADH, E_3 (or FAD), and reconstituted SpnQ. The concentration of each was the same as in the reaction containing the as-isolated native SpnQ described previously. The reaction was analyzed in the same way as were the other SpnQ activity assays.

Determining the Effect of SAM on SpnQ/E_3 Reaction. A total amount of 150 μM SAM was included in the reactions with E_3 or FAD as the potential electron transfer intermediary while maintaining the concentrations of other components unchanged. The reaction was analyzed by HPLC using the same elution program as described for other SpnQ activity assays.

Determining the Effect of the Electron Transfer Protein Pair Flavodoxin/Flavodoxin Reductase on SpnQ Activity - Condition C. Flavodoxin (FLD) and flavodoxin reductase (FNR) are efficient electron transfer proteins that are involved in many biological reactions. Therefore, the activity of FLD/FNR enzyme pair in the SpnQ catalyzed reaction was examined. A typical assay mixture of 50 μL contained 0.70 mM of SpnN product, 0.15 mM PMP, 30 μM flavodoxin, 30 μM flavodoxin reductase, 2.0 mM NADPH, 30 μM SpnQ, and 50 mM potassium phosphate (pH 7.5). A suitable amount of reaction mixture was withdrawn at appropriate time intervals, treated in the same way as described for other SpnQ activity assays, and subjected to HPLC analysis. The same elution program as has been described for other SpnQ activity assays was used to separate the compounds in the reaction mixture.

Determining the Effect of Electron Transfer Protein Pair Swamping on SpnQ Activity – Condition C. To examine the enzymatic activity of all possible combination of electron transfer enzymes, the two pairs of enzymes, ferredoxin/ferredoxin reductase and flavodoxin/flavodoxin reductase were interchanged to produce two more enzyme pairs: ferredoxin/flavodoxin reductase and flavodoxin/ferredoxin reductase. The two new enzyme pairs were individually included in the SpnQ activity assays as described

previously. All the enzyme concentrations were maintained at 30 μM. The electron transfer activities of FLD/FNR, ferredoxin/ferredoxin reductase, FLD/ferredoxin reductase, and ferredoxin/FNR in the SpnQ reaction were compared by their respective relative reaction rates.

Determining the Effect of Including Only One Member of Each Electron Transfer Protein Pair on SpnQ Activity – Condition C. In separate reactions, the enzymes flavodoxin, flavodoxin reductase, ferredoxin, and ferredoxin reductase were individually examined for their electron transfer capabilities. Each enzyme was included separately in an assay mixture as was described for previous SpnQ activity assays. The enzyme was removed by Centricon YM-10 and the filtrate was injected to HPLC to analyze substrate conversion. The same elution program as described previously for other SpnQ activity assays was used.

Regeneration of NADPH – Condition C (Figure 3-4). To facilitate the

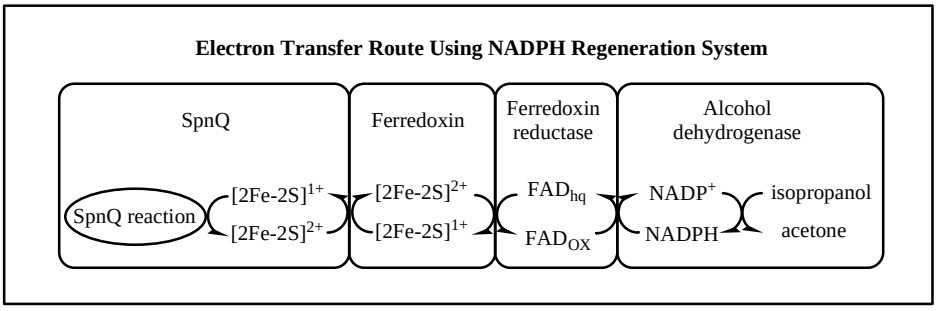


Figure 3-4 . Electrons are provided by isopropanol and transfered to the SpnQ reaction, passing alcohol dehydrogenase, ferredoxin reductase, ferredoxin and reaching SpnQ.

purification of the C-3 deoxygenation product from the reaction mixture, low concentrations of NADPH are desirable since NADPH and $NADP^+$, the oxidized product of NADPH, are the major compounds that need to be removed. To decrease the amount of NADPH used in the reaction, an NADPH regeneration system was employed. A typical assay mixture of 50 μL contained 3.0 mM of SpnN product, 0.15 mM PMP, 30 μM ferredoxin, 30 μM ferredoxin reductase, 60 mM of isopropanol, 0.25 mM of $NADP^+$, 20 μM of alcohol dehydrogenase, 30 μM of SpnQ, and 50 mM potassium phosphate (pH 7.5). The reaction was carried out at 37 °C. A suitable amount of reaction mixture was withdrawn at appropriate time intervals, was treated in the same way as described in the SpnQ activity assay, and was subjected to HPLC analysis. The same elution program as was described for the SpnQ activity assays was used to separate the compounds in the reaction mixture.

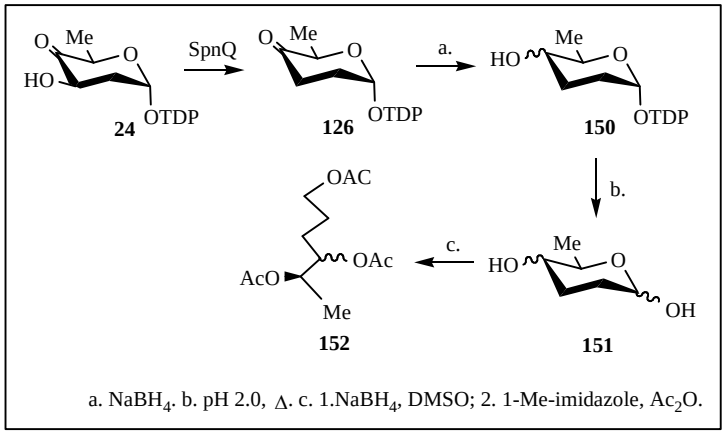
SpnQ-Catalyzed Deprotonation from C-4' of the PMP-Substrate Complex. (4*R*)-[4'-^{3}H]PMP was prepared by Jenefer Alam of our laboratory. The 50 μL reaction mixture contained 0.7 mM SpnN product, 0.15 mM (4*R*)-[4'-^{3}H]PMP, 30 μM SpnQ and 50 mM potassium phosphate buffer (pH 7.5). The reaction was incubated at room temperature overnight before activated charcoal (10% solution, 200 μL) was added. The resulting mixture was mixed vigorously on a vortex for 1 min and centrifuged for 10 min at 13,000 rpm to precipitate the charcoal. A total amount of 100 μL of the supernatant was added to 5 mL of scintillation fluid and analyzed by scintillation counter. Controls were run in parallel in which either the substrate, enzyme, or PMP was omitted in the reaction mixture.

NADPH Concentration Effect on Substrate Conversion – Condition C. NADPH is the reducing cosubstrate in the SpnQ-catalyzed reaction. To study the effect of NADPH concentration on substrate conversion, different concentrations of NADPH: 0.7 mM, 0.84 mM, 1.04 mM, 1.4 mM, 1.76 mM, and 2.1 mM were used in the reactions while the concentration of the substrate was maintained at 0.7 mM. Other components and their concentrations were the same as for the SpnQ activity assay. The percent conversions of the substrate were calculated.

Production and Characterization of the SpnQ Product – Condition C. A preparative-scale reaction contained 15.2 μmol SpnN product, 0.75 μmol PMP, 0.94 μmol $NADP^+$, 0.3 mmol isopropanol, 8.5 mg NADP-dependent alcohol dehydrogenase, 1.8 mg ferredoxin, 0.2 mg ferredoxin reductase, and 6.0 mg SpnQ in 3.6 mL 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of $NADP^+$ and was incubated at 37 °C. The progress of the reaction was monitored by HPLC analysis as described previously. The reaction was almost complete in 2 h. The enzymes in the reaction mixture were removed by filtering through Centricon YM-10 by microcentrifugation. The filtrate was subjected to gel filtration chromatography (Bio-Rad P2, extra fine, 2.5 x 150 cm) equilibrated with 25 mM NH_4HCO_3. The column was eluted with the same buffer at a flow rate of 12 mL/h at 4 °C. Fractions were assayed spectrophotometrically from 250 nm to 400 nm. The fractions that displayed a single absorption maximum at 267 nm were lyophilized individually and analyzed by NMR. Spectra data for the SpnQ product: ^1H NMR (D_2O, mixture of keto and hydrated keto form): δ 0.87, 0.97 (3H, d, J = 6.5 Hz, 5-Me), 1.64 (3H, s, 5’’-Me), 1.54-2.44 (4H, m, 2-

H, 3-H), 2.02-2.13 (2H, m, 2'-H), 3.84 (1H, q, *J* = 6.5 Hz, 5-H), 3.87-3.91 (3H, m, 4'-H, 5'-H), 4.33-4.37 (1H, m, 3'-H), 5.28 (1H, d, *J* = 6.5 Hz, 1-H), 6.08 (1H, t, *J* = 7.0 Hz, 1'-H), 7.50 (1H, s, 6''-H); ^{13}C NMR (D_2O, hydrated form): δ 11.7, 11.8, 28.0, 30.3, 38.6, 70.5, 71.0, 84.9, 85.3, 91.4, 93.7, 93.8, 111.8, 137.4, 151.8, 166.7; ^{31}P NMR (D_2O): δ –10.7 (d, *J* = 21Hz), -12.5 (d, *J* = 21Hz). High-resolution FAB- MS calc for $C_{16}H_{23}N_2O_{13}P_2$ $(M-H)^+$ 513.0675, found 513.0692.

GC/MS of SpnQ Product Derivative. A 800 μL reaction contained 0.84 μmol of the SpnN product, 0.09 μmol PMP, 4.8 μmol NADPH, 0.12 mg ferredoxin, 0.03 mg ferredoxin reductase, 0.4 mg SpnQ in 50 mM potassium phosphate buffer (pH 7.5). The reaction was allowed to proceed at 24 °C for 4 h and then was treated according to Scheme 3-1. $NaBH_4$ (0.38 mg, 0.01 mmol) was added to the reaction mixture. The resulting solution was incubated at room temperature for 30 min before



Scheme 3-1. SpnQ product derivatization for GC/MS analysis.

being acidified to pH 2.0 with 5 M HCl. The solution was boiled for 10 min and centrifuged to remove the enzyme precipitate. The supernatant was then neutralized with 5 M KOH to pH 7.0. The solution was lyophilized to dryness and the solid was redissolved in 0.6 mL of 0.5 M NH_4OH. The sample was mixed with 1.0 mL of DMSO to which $NaBH_4$ (21 mg, 0.55 mmol) had been added. The reaction was stirred at 40 °C for 90 min before being slowly quenched with 0.1 mL of glacial acetic acid. 1-Methyl-imidazole (2 mg, 25 μmol) and 6.0 mL of acetic anhydride were added, and the resulting mixture was stirred at room temperature for 1 h. The reaction was then quenched with 5.0 mL of methanol at 0 °C. The products were extracted with chloroform and the combined organic extracts were washed twice with water, dried over anhydrous sodium sulfate, filtered, and concentrated. The sample was then subjected to GC/MS analysis. A control reaction was run in parallel in which the SpnQ protein was omitted from the reaction.

3. RESULTS AND DISCUSSION

Cloning and Expression of SpnQ Protein. Cloning of *spnQ* into pET24b(+) vector gave two constructs that produced either native SpnQ or *C*-terminal His-tagged SpnQ, respectively, depending on the presence or absence of the natural stop codon of the gene. Both proteins were partially soluble in the expression cells and accounted for about 10% of the total proteins in the crude extract.

Purification and Characterization of His-tagged SpnQ Protein (Figure 3-5). The His-tagged enzyme was purified to near homogeneity by a Ni-NTA chromatography

(Figure 3-5). A total of 40 mg of homogeneous protein was obtained from a 6 L culture. The estimated subunit molecular weight of 51 kDa revealed by SDS-PAGE is consistent with the molecular weight 51446 Da, deduced from the translated amino acid sequence of SpnQ (including His_6-tag).

Purification and Characterization of Native SpnQ Protein (Figure 3-5). Various types of chromatography were tried in the purification of native SpnQ. The protein has poor binding affinity to phenyl-Sepharose resin and binds weakly to Mono Q resin. The protein was finally purified to near homogeneity by sequential ammonium sulfate precipitation, DEAE chromatography, Hydroxyapatite chromatography, and S-200 chromatography. A total of 30 mg of light red-brown protein was obtained from a 6 L culture. Including 15% glycerol was crucial in minimizing protein precipitation during the purification process. Its subunit molecular weight of 50 kDa, as revealed by SDS-PAGE,

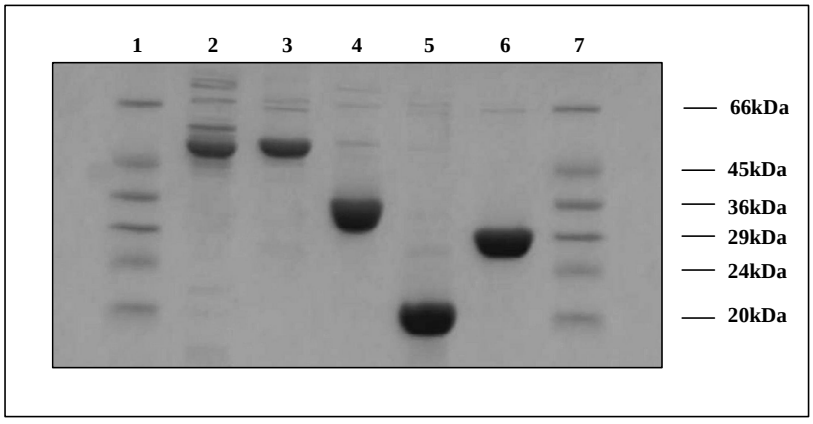


Figure 3-5. SDS-polyacrylamide gel electrophoresis of proteins obtained in Chapter 3. Lanes 1 and 7, molecular markers; lane 2, native SpnQ protein; lane 3, *C*-terminal His-tagged SpnQ; lane 4, E_3 protein ; lane 5, flavodoxin ; lane 6, flavodoxin reductase.

correlates well to the predicted value of 50381 Da based on the translated peptide sequence. The molecular weight of native SpnQ was estimated to be 109.4 kDa by gel filtration chromatography, suggesting that the protein exists as a homodimer in solution. The sequence of the first 15 amino acid residues at the *N*-terminus of SpnQ was determined to be M-Q-S-R-K-T-R-A-L-G-K-G-R-A-R, which matches exactly with the translated *spnQ* sequence.

Preparation of Apo-SpnQ. The iron-sulfur center of the as-purified SpnQ was completely depleted within 1.5 h after the addition of dithionite and EDTA as judged by the complete loss of its light red-brown color. The electronic absorption spectrum of Apo-SpnQ shows no absorbance above 300 nm.

SpnQ Reconstitution. Both apo-SpnQ and as-isolated native SpnQ were reconstituted in an anaerobic environment following the same protocol. Both reconstituted proteins have dark green-black color which is common to many enzymes whose iron-sulfur centers are reconstituted.[178,179] Interestingly, the enzyme reconstituted from Apo-SpnQ did not show any enzymatic activity in any of the three assay conditions: condition A, B or C. This observation hints that SpnQ might undergo some irreversible conformational change when the iron-sulfur center is completely removed. Subsequent repletion with iron and sulfur component cannot restore the correct enzyme conformation. The enzymatic activity of the enzyme reconstituted from the as-isolated SpnQ will be discussed later.

UV-visible Spectrum of SpnQs. As shown in Figure 3-6, the spectrum of native SpnQ has bands centered around 330 nm, and 460 nm with a broad shoulder centered around 550 nm.[159,160] His_6-tagged SpnQ has absorption maxima at the same wavelengths but the absorbance is less significant considering the protein concentration. In view of the literature value of ε_{460} in the range of 6000-10000 $M^{-1}cm^{-1}$ for the [2Fe-2S] center and the enzyme concentrations measured by UV absorbance at 280 nm, the [2Fe-2S] content of His_6-tagged SpnQ is between 0.25 and 0.42 clusters per subunit while that of native SpnQ is between 0.36 and 0.61 clusters per subunit.[161] The anaerobically reconstituted SpnQ also has absorption maxima at 330 nm, 460 nm and 550 nm, but the absorbance at the absorption maxima is significantly enhanced.

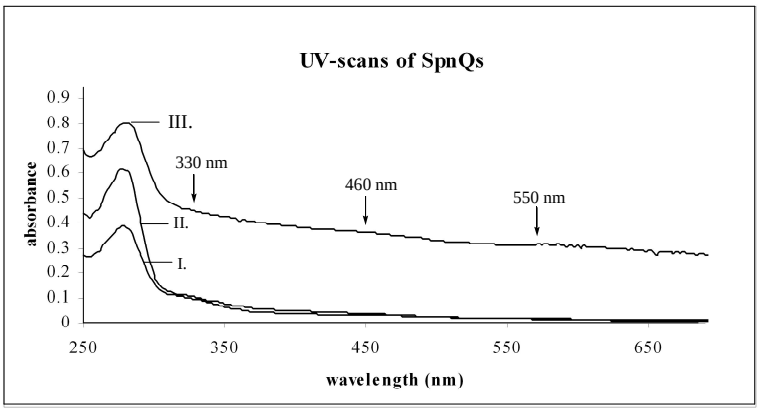


Figure 3-6 . UV-scan spectra of SpnQs. (I). the as-isolated native SpnQ protein. (II). *C*-terminal His_6-tagged SpnQ. (III). SpnQ reconstituted from the as-isolated native protein.

Iron Content of SpnQ Protein. Compared to the iron content standard curve (Figure 3-7), native SpnQ contains 1.2 irons per polypeptide. This value is relatively

consistent with 0.36~0.61 [2Fe-2S] clusters per subunit as judged by ε_{460}. The anaerobically reconstituted SpnQ contains 4.1 iron per polypeptide, suggesting that there are two [2Fe-2S] clusters per polypeptide in SpnQ.

Purification of E_3, FLD and FNR (Figure 3-5). Purification of E_3 was done according to an established procedure.[57,62] After ammonium sulfate precipitation, DEAE chromatography, and phenyl-Sepharose chromatography, 120 mg of nearly homogeneous E_3 was obtained from a 6 L culture. FLD and FNR were each purified to near homogeneity by a single Ni-NTA chromatography step. About 15 mg of each FLD and FNR was obtained from 2 L of culture. The electron transfer activity of E_3 was demonstrated by the well-established DCPIP assay,[180,181] while the enzymatic activity of FLD/FNR was demonstrated by the ferricyanide assay.[182]

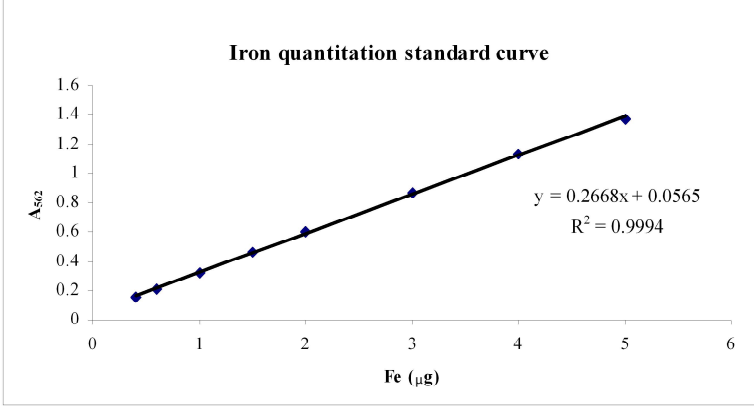


Figure 3- 7. Iron quantitation standard curve.

Isolation and Characterization of the SpnQ Product-Condition A (Figure 3-8). When the SpnQ reaction (condition A) was analyzed by HPLC, the formation of a new compound was observed accompanied by the simultaneous consumption of the substrate. The new compound was originally thought to be the C-3 deoxygenated product, TDP-2,3,6-trideoxy-4-keto-glucose (**126**). However, instead of having a *M.W.* of 514 of TDP-2,3,6-trideoxy-4-keto-glucose (**126**), the *M.W.* of this compound is 531 by ESI-MS analysis. Interestingly, this value matches well with the *M.W.* of TDP-4-amino-

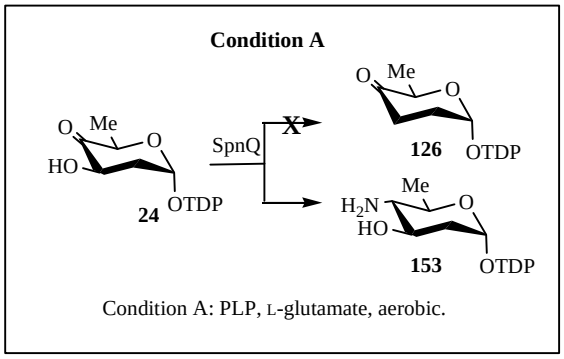


Figure 3-8 . SpnQ catalyzed reaction under condition A.

2,4,6-trideoxy-glucose (**153**). It has been mentioned previously that the sequence of SpnQ is similar to a number of aminotransferases. In view of the components used in the assay mixture other than the SpnN product, PLP and L-glutamate, the only possible rationalization is that SpnQ under condition A catalyzes the C-4 aminotransfer reaction. For the TDP-hexoses that undergo transamination reactions, the substrate and the product are generally well separated by HPLC using a Dionex column. Thus, it is not surprising

to observe that the retention time of the new compound from the SpnQ reaction under condition A is quite different from that of the SpnN product. In the presence of PLP and L-glutamate, ColD catalyzes C-3 deoxygenation, while under the same conditions, SpnQ shows C-4 aminotransferase activity. Thus, the mechanism of SpnQ-catalyzed C-3 deoxygenation must be different from that of the ColD catalyzed reaction (mechanism A). The aminotransfer reaction could be a side reaction that SpnQ catalyzes or it could be part of a complicated C-3 deoxygenation reaction. It is interesting to note that this aminotransfer reaction is quite similar to DesI catalyzed reaction except that the substrate is TDP-2,6-dideoxy-4-keto-glucose (**24**) instead of TDP-6-deoxy-4-keto-glucose (**3**).

Amino Donors - Condition A. In the SpnQ catalyzed C-4 aminotransfer reaction, L-glutamate was found to be an efficient amino donor while L-aspartate could hardly

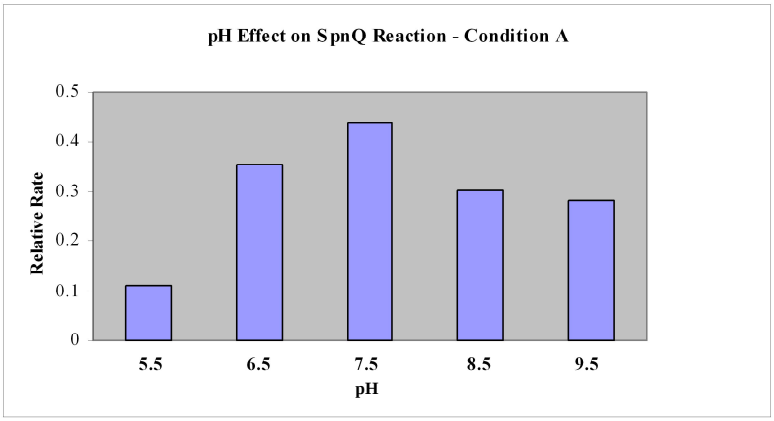


Figure 3-9 . Effect of pH on the relative rate of SpnQ reaction under condition A.

function as an amino donor. When L-aspartate was used, SpnQ displayed 5% of the activity observed with L-glutamate.

Optimal pH for SpnQ - Condition A. As shown in Figure 3-9, SpnQ has an optimal aminotransferase activity at around pH 7.5. Under acidic and alkaline conditions, the substrate, the SpnN product, was unstable and degraded. The degradation was more severe under the alkaline conditions than under the acidic conditions.

Temperature Effect for SpnQ - Condition A. The optimal aminotransferase activity of SpnQ was at 37 °C. With decreasing temperatures, a sharp reduction in activity was observed. At 24 °C, almost no substrate conversion was observed.

Effect of DTT. Ornithine decarboxylase is a PLP-dependent enzyme. The catalytic activity of ornithine decarboxylase was stimulated in the presence of DTT.[162] The effect of DTT on the aminotransferase activity of SpnQ was examined. At the level of 5 mM, which has significantly increased ornithine decarboxylase activity,[162] DTT failed to show any enhancing effect on SpnQ activity.

Effect of SAM. Cystathionine β-synthase (CBS) catalyzes the PLP-dependent β-replacement reaction in which the thiolate of L-homocysteine replaces the hydroxyl group of L-serine. Human CBS has a complex domain structure, and its regulatory mechanism is complicated. The *N*-terminal domain of CBS is a heme-binding domain. The allosteric activator SAM increases CBS activity about 3-fold and likely binds to the *C*-terminal regulatory domain.[163,164] Since SpnQ is also a coenzyme B_6-dependent enzyme and has an iron-sulfur center which might bind SAM, the possibility of SAM as an allosteric activator was examined. The results show that the presence of SAM does not have any stimulatory or inhibitory effect on the aminotransferase activity of SpnQ.

Effect of SpnQ Reconstitution on SpnQ Activity – Condition A. The results from iron quantitation studies and UV analysis have shown that the iron-sulfur content of the reconstituted enzyme was increased significantly. To examine any possible connection between the iron-sulfur content and the aminotransferase activity of SpnQ, the anaerobically reconstituted enzyme was used to conduct the activity assay. The reconstituted enzyme did not show any change in the level of aminotransferase activity compared with the as-isolated enzyme. It can thus be inferred that the iron-sulfur cluster does not participate in the aminotransfer reaction.

Effect of SpnQ His_6-tag on Aminotransferase Activity – Condition A. Interestingly, the His_6-tagged enzyme did not show any aminotransferase activity. The His_6-tag probably prevents the correct folding of the enzyme. Thus, His_6-tagged enzyme was not used in further enzyme assays.

Reduction by Dithionite -Condition B (Figure 3-10). Dithionite is a strong

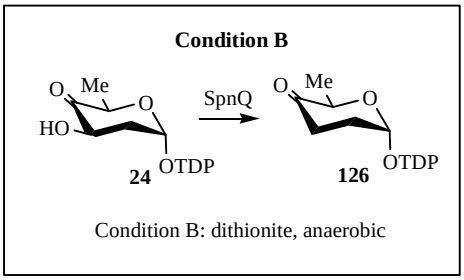


Figure 3-10 . SpnQ catalyzed reaction under condition B.

reducing agent and is able to provide electrons to reduction reactions under anaerobic conditions. When the reaction is exposed to air, because of the high oxidation potential

of O_2, O_2 becomes the direct electron acceptor preventing other oxidants from being reduced by dithionite. Thus, the assays with dithionite as the only reducing agent were conducted anaerobically. SpnQ was able to catalyze the C-3 deoxygenation to produce TDP-2,3,6-trideoxy-4-keto-glucose (**126**). The conclusion that TDP-2,3,6-trideoxy-4-keto-glucose was indeed being formed came from two pieces of evidence: first, an ^{1}H-NMR spectrum was obtained from a preparative scale reaction. Although the spectrum was contaminated with peaks from impurities, the peaks of TDP-2,3,6-trideoxy-4-keto-glucose were all present; second, a new compound formed accompanied by the decrease of the substrate as determined by HPLC analysis, and the new compound had the same retention time as TDP-2,3,6-trideoxy-4-keto-glucose analyzed by NMR. Dithionite was also able to provide single electrons to PMP-$\Delta^{3,4}$-glucoseen intermediate in the E_1-catalyzed C-3 deoxygenation in the biosynthesis of CDP-ascarylose when E_3 and NADH were absent from the reaction mixture, as was seen in these assays of SpnQ.

Concentration Effect of Dithionite - Condition B. Interestingly, it was found that when the concentration of dithionite used in the reaction was comparable to that of the substrate, C-3 deoxygenation was inhibited. When the concentration of dithionite was higher than 3.0 mM while the concentration of the substrate is 1.0 mM, the substrate conversion was lower than 10%. The low conversion rate made product purification difficult and was the reason for the contaminated ^{1}H-NMR spectrum mentioned above. It is speculated that when the concentration of dithionite was increased, the reduction potential of dithionite was correspondingly enhanced. This may cause damage to the integrity of the [2Fe-2S] center in SpnQ since iron-sulfur centers have multiple forms,

such as [4Fe-4S] and [3Fe-4S] and these forms can be interconverted.[183,184] However, evidence for an explanation of this observed inhibition by high concentrations of dithionite must awaits more detailed studies.

Effect of SAM on SpnQ Activity - Condition B. DesII, the C-4 deoxygenation enzyme in the TDP-desosamine biosynthesis, is SAM radical dependent. The *in vitro* enzymatic activity of DesII was demonstrated when dithionite was used as the reducing agent in the reaction. As has been described previously, it is possible that the SpnQ-catalyzed C-3 deoxygenation has a similar mechanism to the DesII-catalyzed C-4 deoxygenation provided that there is a nonenzymatic tautomerization between TDP-2,6-dideoxy-4-keto-glucose (**24,** Figure 3-2) and TDP-2,6-dideoxy-3-keto-glucose (**136**, Figure 3-2). Although SAM was not included in the SpnQ reaction mixture (condition B), the possibility that there was already SAM present that was tightly bound to SpnQ and functioned in the C-3 deoxygenation, could not be eliminated since SAM was reported to bind close to the iron-sulfur center in SAM radical dependent enzymes.[132,185] If that is the case, then increasing the concentration of SAM is expected to facilitate the reaction and increase the reaction rate. Thus, a catalytic amount of SAM and an amount comparable to that of the substrate were each examined for their effects on the reaction rate. At either concentration, SAM did not show any enhancing effect. Moreover, in one SpnQ reconstitution experiment, after the addition of PLP and PMP, SAM of the same concentration as PLP and PMP was added to the incubation mixture before $Fe(NH_4)_2(SO_4)_2$ and Na_2S were added. Enzyme reconstituted in this way was used in the activity assays (condition B). The relative reaction rate of C-3 deoxygenation exhibited

no change compared with the assays using as-isolated SpnQ or SpnQ reconstituted without SAM. Combining all this information, it can be concluded that SAM does not play a role in the SpnQ catalyzed C-3 deoxygenation. This eliminated the possibility that the SpnQ-catalyzed reaction is SAM radical dependent.

Isolation and Characterization of the SpnQ Product - Condition C (Figure 3-11). Isolation of the C-3 deoxygenation product, TDP-2,3,6-trideoxy-4-keto-glucose (**126**) , had been problematic mainly due to observed substrate inhibition. One NADPH molecule provides two electrons both of which are used in the reduction of the PMP-substrate complex. Thus the ratio between NADPH and the substrate should be 1:1 for the overall reaction. When the concentration of the substrate was increased, the

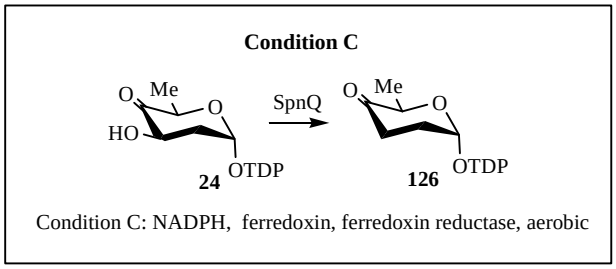


Figure 3- 11. SpnQ catalyzed reaction under condition C.

concentration of NADPH had to be correspondingly increased. However, it was observed that substrate conversion became low (10%) when both the substrate and NADPH concentrations were high (3 mM). In order to circumvent this problem, NADPH was regenerated *in situ* by the catalysis of alcohol dehydrogenase with isopropanol as the cosubstrate. The amount of isopropanol added in the reaction mixture

was 20 times the amount of the SpnN product in order to drive the reaction to completion. The amount of NADP in the reaction mixture was just enough to initiate the reaction and the by-product acetone was easily removed. Clean NMR spectra of the SpnQ product were obtained from this method. Spectroscopic analysis confirmed that the product was the expected C-3 deoxygenation product (Figure 3-12). Due to the removal of the hydroxyl group at C-3, the chemical shift of the C-3 protons moved upfield. As a result, the overlapping between the C-3 and C-5 protons was eliminated. The C-5 proton was left as a simple quartet. Meanwhile, the spectrum around the C-2 protons became complicated due to overlapping with the C-3 protons. The enzymatic conversion was almost quantitative.

Ferredoxin reductases are ubiquitous flavoenzymes that deliver electron equivalents from NADPH to low potential one-electron donors, such as ferredoxin, which are used in redox-based metabolism in plastids, mitochondria and bacteria.[186,187] Ferredoxin reductases constitute a family of hydrophilic, monomeric enzymes that contain noncovalently bound FAD as a prosthetic group, while ferredoxins are [2Fe-2S] containing enzymes.[188] When NADPH is the electron source, two electron equivalents are transferred to cofactor FAD of ferredoxin reductase. The cofactor FAD, which is capable of transferring either a single electron or two electrons, functions as the $2e^-/1e^-$ switch and delivers the electrons one at a time to the [2Fe-2S] center, an obligatory one-electron redox coenzyme of ferredoxin. Thus the combination of ferredoxin and ferredoxin reductase has both FAD and a [2Fe-2S] center, and functions in the same way

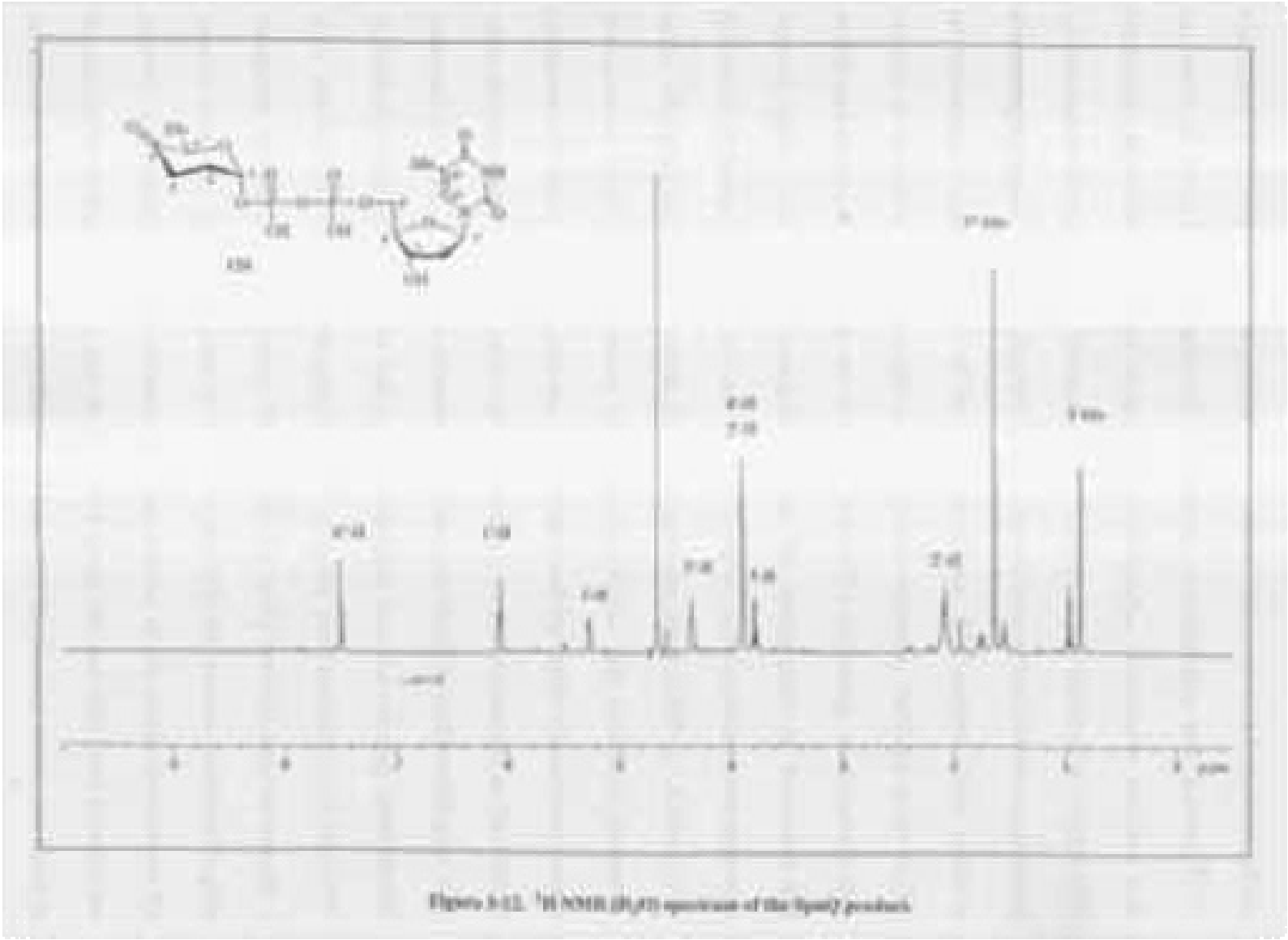


Figure 3-12. ^{1}H NMR (D$_2$O) spectrum of the lipid product.

as E_3. The successful isolation of the C-3 deoxygenation product TDP-2,3,6-trideoxy-4-keto-glucose (**126**) in the presence of SpnQ, ferredoxin, and ferredoxin reductase, combined with the fact that the sequences of E_1 and SpnQ are highly homologous to each other, points out that the SpnQ-catalyzed reaction has a similar mechanism as the reaction catalyzed by E_1/E_3. Thus the mechanism C proposed previously is preferred (Figure 3-3). In the E_1/E_3 reaction, the enzyme that transfers electrons from the reducing agent NADH is E_3 which is encoded by a gene physically close to the gene encoding E_1 in the CDP-ascarylose biosynthetic gene cluster. In contrast, the SpnQ reaction does not employ a specific reductase enzyme encoded by a gene in the TDP-forosamine biosynthetic gene cluster. As evidenced by the demonstrated dependence of the SpnQ reaction on the ferredoxin/ferredoxin reductase pair *in vitro*, and given the fact that these proteins are ubiquitous in the cell, this pair is a likely candidate for the electron source for SpnQ *in vivo*. The E_1/E_3 reaction and the SpnQ reaction are also different in their use of reducing agents. NADPH has been shown to be the favored electron source for ferredoxin/ferredoxin reductase, and thus is used in the SpnQ reactions. Conversely, NADH is the established electron source for E_3 and is used in the E_1/E_3 reaction.

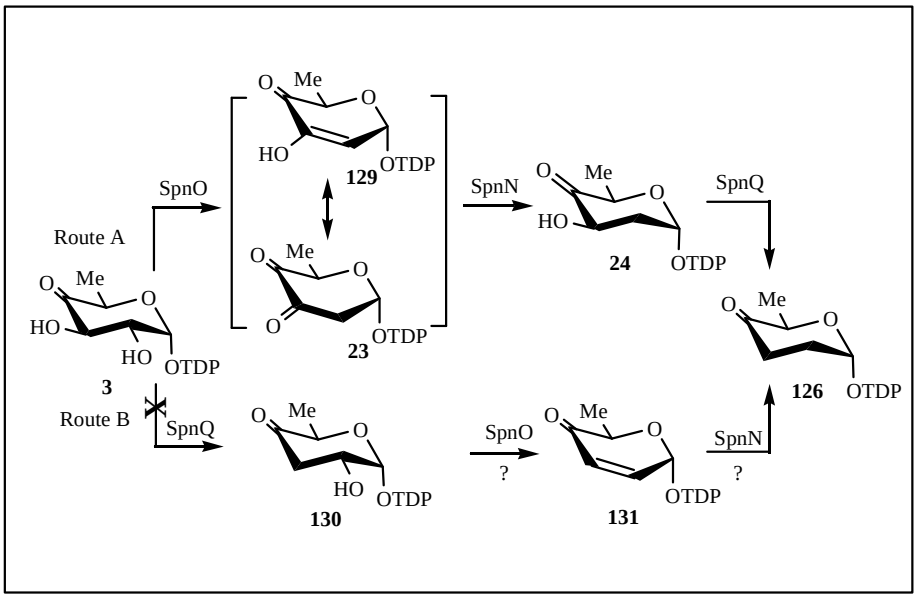
E_1/E_3 Reaction with the SpnN Product as Substrate – Condition C. The sequences of SpnQ and E_1 are highly similar to each other (49% identity). However, a gene encoding an E_3 homolog has not been found in the biosynthetic gene cluster of TDP-forosamine. In order to examine whether the SpnQ-catalyzed reaction is similar to the E_1/E_3 catalyzed reaction, enzymes E_1 and E_3 were used in an assay containing the SpnN product as the substrate (condition C). No turnover of the substrate was observed.

However, failure to observe the C-3 deoxygenation activity of E_1/E_3 may be because the structure of the SpnN product is different from the authentic substrate of E_1/E_3. In E_1/E_3 catalyzed C-3 deoxygenation in CDP-ascarylose biosynthetic pathway, the substrate is CDP-6-deoxy-4-keto-D-glucose, while in this reaction, the substrate is TDP-2,6-dideoxy-4-keto-D-glucose (**24**). E_1/E_3 may not be able to recognize a TDP-sugar as a substrate. The lack of a hydroxyl group at C-2 may also prevent the SpnN product from being accepted by E_1/E_3.

E_3 and FAD as Possible Electron Transfer Intermediaries in SpnQ-Catalyzed Reaction – Condition C. To examine whether an E_3 homolog is required in the SpnQ catalyzed reaction, E_3 was included in the assay mixture. No turnover of the substrate was observed. Thus the SpnQ/E_3 pair could not carry out C-3 deoxygenation in spite of the sequence similarity between E_1 and SpnQ. It was next speculated that SpnQ may only need an electron transfer cofactor such as FAD instead of an enzyme to transfer electrons from NADH to the substrate complex. Electron transfer from NADH to FAD has been documented in the catalysis of methane monooxygenase.[157,158] However, when SpnQ was incubated with the SpnN product, NADH, and FAD, no substrate turnover was observed, thus disproving this possibility.

Effect of SpnQ Reconstitution. To provide compelling evidence that the SpnQ-catalyzed C-3 deoxygenation reaction does not need an enzyme homologous to E_3 or a cofactor, like FAD, as an electron intermediary, the SpnQ enzyme was anaerobically reconstituted to ensure that it contains all necessary cofactors. With that done, failure to observe C-3 deoxygenation can be conclusively ascribed to the lack of a requirement for

FAD and present persuasive evidence that an E_3 homolog is not required. Thus, the SpnQ/E_3 (FAD) reactions described previously were run again using reconstituted SpnQ, which was considered to be in its active form with all the cofactors present. Even under these conditions, there was still no substrate conversion. Therefore, the inability of SpnQ/E_3 (FAD) to catalyze C-3 deoxygenation is not due to the incomplete iron-sulfur center and PMP or PLP contents in the as-isolated SpnQ. Thus, both E_3 and FAD have been eliminated as NADH-dependent electron transfer intermediaries in the SpnQ-catalyzed reaction. It can also be inferred from these results that it is unlikely that there is



Scheme 3-2 . Between the two possilbe routes to obtain 126, Route B, 3-deoxygenation followed by 2-deoxygenation was eliminated. SpnQ can only recognize the SpnN product 24.

an E_3 homolog in the TDP-forosamine biosynthetic gene cluster which aids in C-3 deoxygenation.

Substrate Specificity of SpnQ –Condition C (Scheme 3-2). When TDP-6-deoxy-4-keto-glucose (**3**), instead of the SpnN product, was used in the assay reactions, the peak of TDP-6-deoxy-4-keto-glucose remained unchanged in HPLC analysis. However, this observation alone could not completely eliminate the possibility that there was a conversion. If the retention time of the putative conversion product, TDP-3,6-dideoxy-4-keto-glucose (**130**) , happened to be the same as that of TDP-6-deoxy-4-keto-glucose (**3**),no change would be observed in the HPLC trace even though conversion had occurred. If this was the case, then adding SpnO, SpnN and more NADPH would have converted TDP-3,6-dideoxy-4-keto-glucose (**130**) to TDP-2,3,6-trideoxy-4-keto-glucose (**126**) which had been shown to have a different retention time from that of TDP-6-deoxy-4-keto-glucose (**3**) under the HPLC conditions used. No peak corresponding to TDP-2,3,6-trideoxy-4-keto-glucose was observed while the peak of TDP-6-deoxy-4-keto-glucose remained largely unchanged in HPLC analysis. The tiny amount of the SpnN product, TDP-2,6-dideoxy-4-keto-glucose (**24**), formed by the action of SpnO and SpnN towards **3,** was likely to be too little to be observed. As was explained in Chapter 2, SpnO shows every low activity *in vitro*. In the UV-absorbance activity assay for SpnO, the decrease of A_{340} corresponding to the formation of the SpnN product was very small. There had been doubt whether the C-2 deoxygenation or the C-3 deoxygenation occurs first in the TDP-forosamine biosynthetic pathway. The result from this experiment clearly indicated that SpnQ could not recognize TDP-6-deoxy-4-keto-glucose as a substrate. So the RfbB product can only be the substrate of the SpnO/SpnN catalyzed C-2 deoxygenation, while the product thus formed becomes the substrate of the SpnQ-

catalyzed C-3 deoxygenation. Thus, C-2 deoxygenation occurs prior to C-3 deoxygenation. This observation will likely hold true in the biosynthesis of any NDP-2,3-dideoxy-hexoses in whose biosynthetic gene cluster proteins homologous to SpnQ and TylX3/SpnO are identified.

Effect of SAM – Condition C. Addition of SAM in either a catalytic amount or in an amount comparable to the substrate did not show any effect on substrate conversion. In the enzyme assays under all three conditions (condition A, B or C), SAM did not show any effect on catalysis, nor did it participate in the reactions by acting as cosubstrate. It can now be concluded that neither the SpnQ-catalyzed C-4 transamination nor the C-3 deoxygenation requires SAM. Thus, the C-3 deoxygenation mechanism must be completely different from the SAM radical dependent mechanism of the DesII-catalyzed C-4 deoxygenation. Therefore, mechanism B proposed earlier was conclusively eliminated.

Electron Transfer Proteins in the SpnQ-Catalyzed Reaction - Condition C (Figure 3-13). The enzyme pair of ferredoxin/ferredoxin reductase has been shown to transfer electrons from NADPH to SpnQ in the SpnQ-catalyzed C-3 deoxygenation. The ubiquitous presence of this enzyme pair *in vivo* implicates that the reductase required in the SpnQ reaction may come from the general cellular pool. However, ferredoxin and ferredoxin reductase are not the only electron transfer proteins in the cell. Flavodoxin and flavodoxin reductase, containing [2Fe-2S] and FAD respectively, can catalyze

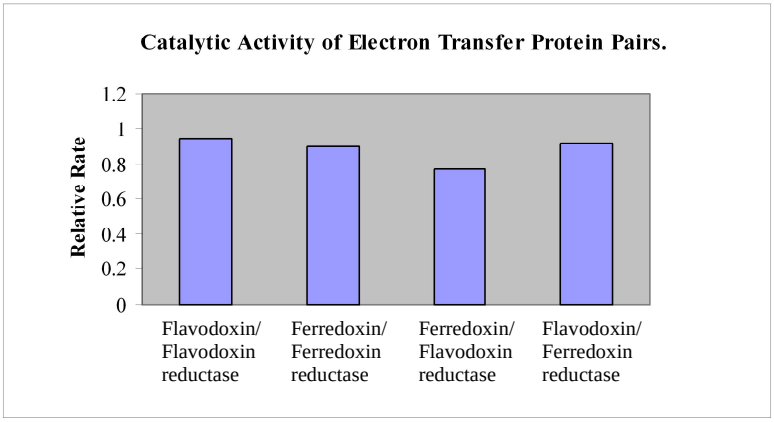


Figure 3-13. Catalytic activities demonstrated by different enzyme pairs.

electron transfers in the same way as ferredoxin and ferredoxin reductase. When the enzyme pair of ferredoxin/ferredoxin reductase in the SpnQ activity assay was replaced with flavodoxin/flavodoxin reductase, the latter enzyme pair showed reductase competency comparable to the former enzyme pair as judged by the substrate conversion rate. The components of ferredoxin/ferredoxin reductase and flavodoxin/flavodoxin reductase were interchanged to make two new enzymes pairs, ferredoxin/flavodoxin reductase and flavodoxin/ferredoxin reductase. The enzyme pair flavodoxin/ferredoxin reductase was almost as competent as the natural enzyme pairs, while ferredoxin/flavodoxin reductase has slightly lower catalytic activity (about 80% of the other three pairs). Any single component, ferredoxin, flavodoxin, ferredoxin reductase or flavodoxin reductase did not show any electron transfer activity, as evidenced by the absence of substrate conversion. Interestingly, the enzyme pair flavodoxin/flavodoxin reductase completely lost its electron transfer activity after several freeze-thaw cycles,

while the enzyme pair ferredoxin/ferredoxin reductase maintained full activity for months. Moreover, the enzyme pair ferredoxin/flavodoxin reductase maintained 40% of its activity while the enzyme pair flavodoxin/ferredoxin reductase maintained 20% of its activity after several freeze-thaw cycles. Thus, flavodoxin and flavodoxin reductase each individually lost some activity by the freeze-thaw treatment. They were still partially active but completely lost the ability to act as a functional pair after degradation.

Regeneration of NADPH. The NADPH regeneration system made of alcohol dehydrogenase, isopropanol, and a catalytic amount of $NADP^+$, is capable of supplying *in situ* generated electron equivalents from NADPH to the electron transfer chain in the SpnQ reaction. The reaction was efficient at 37 °C but did not occur at all at 24 °C. The temperature dependence of the reaction is not due to SpnQ, but rather to the purchased alcohol dehydrogenase which is recommended to be used at 37° C. Both the substrate and product of the SpnQ reaction were shown to be relatively stable for 2 h at 37 °C.

Stereochemistry of SpnQ-Catalyzed Deprotonation of the PMP-substrate Complex (Figure 3-14). In the E_1/E_3 reaction, the reaction steps from the Schiff base formation (**67**, Scheme 1-10, Chapter 1) to the CDP-6-deoxy-$\Delta^{3,4}$-glucoseen formation (**69**, Scheme 1-10, Chapter 1) are reversible and it is the pro-S proton of C-4' of PMP that is removed and reinstalled.[53,63,64] Since the SpnQ/ferredoxin/ferredoxin reductase reaction is mechanistically similar to the E_1/E_3 reaction, the reaction steps from the Schiff base formation to the TDP-2,6-dideoxy-$\Delta^{3,4}$-glucoseen (**145**) formation should also be

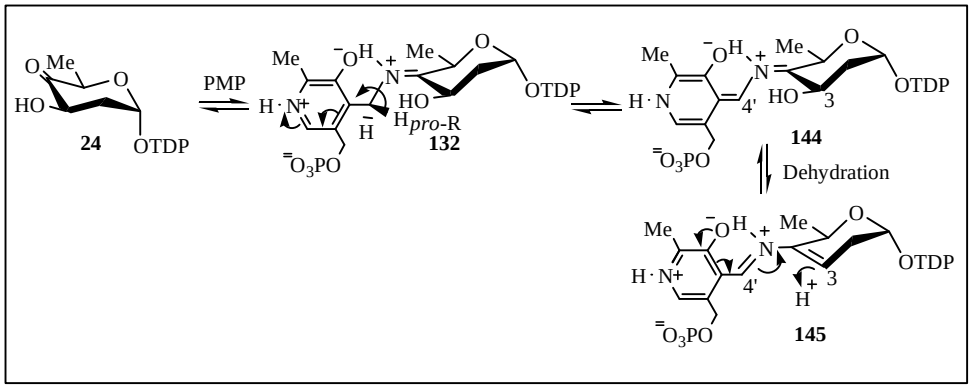


Figure 3-14. Stereospecificity of SpnQ catalyzed deprotonation of PMP-substrate complex.

reversible in the presence of SpnQ. When (4'R)-[4'-^{3}H] PMP was incubated with the substrate and SpnQ, twice the amount of radioactivity was released into the reaction buffer compared with control reactions in which either the substrate, the enzyme or both the substrate and the enzyme was/were omitted from the incubation mixture. This result proves that the deprotonation and reprotonation of the pro-R C-4' proton of PMP are indeed reversible in the SpnQ reaction. However, the evidence obtained so far indicates that the stereochemistry of the deprotonation of PMP is different from that in the E_1/E_3 reaction.

Effect of NADPH Concentration – Condition C. When the concentration of the substrate was 0.7 mM, one equivalent of NADPH could efficiently drive the reaction to completion within 3 h. The conversion of the substrate was also complete at concentration of NADPH as high as 3-fold molar excess relative to the concentration of the substrate. Thus, NADPH did not show any inhibitory effect. However, when the concentrations of both the substrate and NADPH were increased to 3.0 mM, the reaction

rate significantly decreased and only 10% of the substrate was converted to product at equilibrium. In addition, by using the NADPH regeneration system described previously, the substrate present at a concentration of 3.0 mM could be quantitatively converted to product in the presence of a catalytic amount of NADPH. The reason why the SpnQ reaction was inhibited when the concentrations of both the substrate and NADPH were high is still not clear and must await detailed studies.

GC/MS of the SpnQ Product Derivative – Condition C. GC/MS analysis of the SpnQ reaction clearly showed the presence of a compound of m/z 260 which is the expected mass of the derivatized SpnQ product. The m/z 260 peak was absent from the GC trace of the reaction without SpnQ.

4. CONCLUSIONS (Figure 3-15)

Extensive studies of the SpnQ protein were described in this chapter. Both the His_6-tagged and the native protein were expressed and purified. Anaerobic reconstitution was conducted on both the as-isolated enzyme and the apo-enzyme. SpnQ was shown to be an aminotransferase when only PLP and L-glutamate were included in the assay mixture. The product of this reaction is TDP-2,6-dideoxy-4-amino-glucose (**153**). SpnQ was also shown to function as an enzyme catalyzing the C-3 deoxygenation of TDP-2,6-dideoxy-4-keto-glucose (**24**) in the presence of NADPH and ferredoxin/ferredoxin reductase. Flavodoxin/flavodoxin reductase was also shown to function effectively as the electron source of the SpnQ reaction *in vitro*. Both enzyme pairs are ubiquitous in the cell, indicating that either pair may be the electron source of the SpnQ reaction *in vivo*.

SpnQ could also catalyze C-3 deoxygenation under anaerobic conditions with dithionite as the electron source.

Mechanistically, SpnQ catalyzes a radical reaction similar to that of E_1/E_3. But unlike E_1/E_3, SpnQ does not use a dedicated NADH-dependent reductase. This explains well why a gene encoding an E_3 homolog has not been identified in the TDP-forosamine biosynthetic gene cluster. These results have unambiguously demonstrated that the mechanism of the SpnQ catalyzed C-3 deoxygenation is completely different from that of the ColD catalyzed C-3 deoxygenation and the DesII catalyzed C-4 deoxygenation.

The fact that SpnQ uses the SpnN product (**24**) as its substrate, but does not recognize the RfbB product (**3**) clearly established that C-3 deoxygenation occurs after the SpnO/SpnN catalyzed C-2 deoxygenation, thus providing an unequivocal proof for the sequence of the early reactions in TDP-forosamine biosynthesis. SpnQ has homologs in many NDP-2,3-dideoxy-hexose biosynthetic gene clusters. Evidence for the mechanism of the SpnQ catalyzed reaction and the explicit demonstration of the substrate specificity of SpnQ not only provides significant mechanistic insights into the formation of 2,3,6-trideoxy-sugars, but also facilitates biosynthetic studies of other highly deoxygenated sugars.

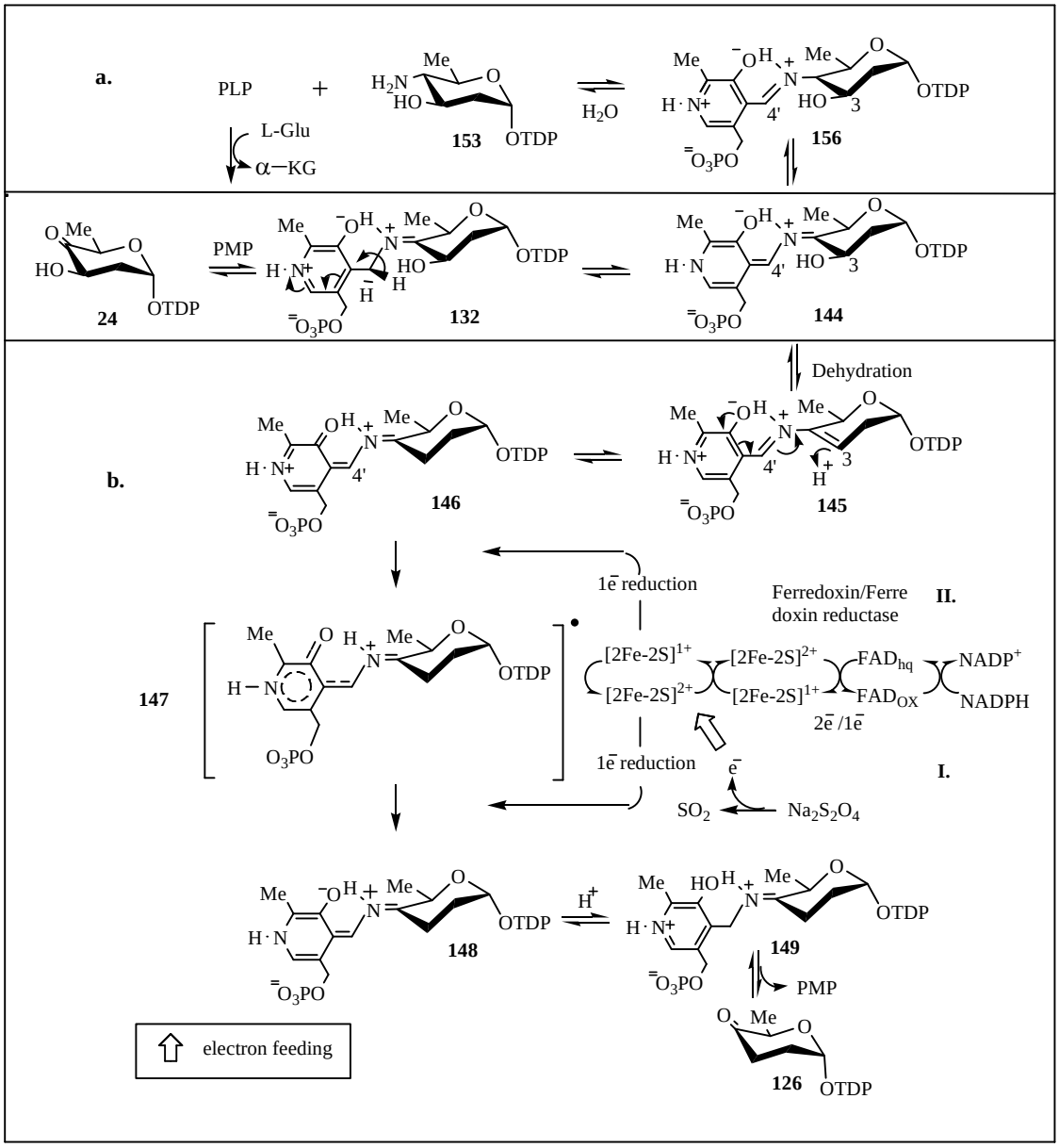


Figure 3-15. Proposed mechanisms of SpnQ catalyzed reactions. (a). SpnQ catalyzes C-4 aminotransfer reaction in the presence of L-glutamate and PLP. (b). SpnQ catalyzes C-3 deoxygenation. The electron can be directly provided by $Na_2S_2O_4$ (I), or can be provided by NADPH and transferred by ferredoxin/ferredoxin reductase (II).

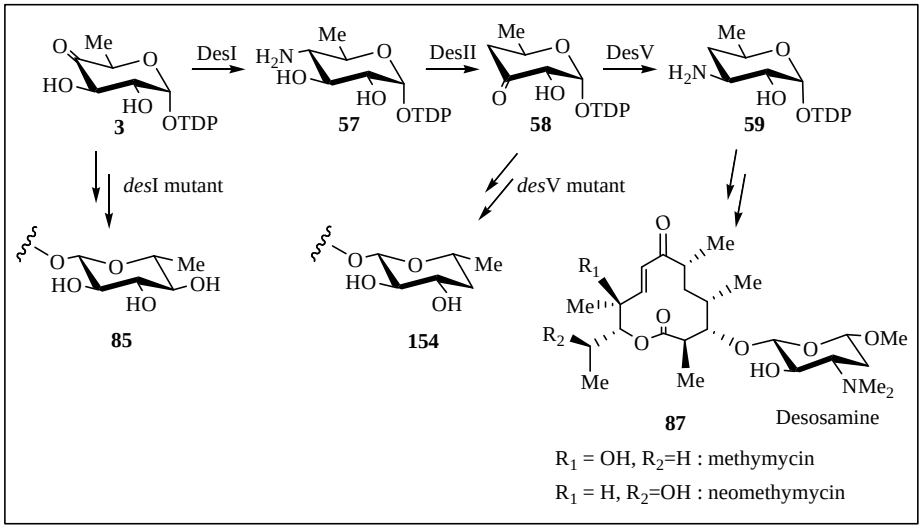
Chapter 4. Characterization of SpnR, the TDP-2,3,6-trideoxy-4- keto-D-glucose 4-aminotransferase and SpnS, the TDP-4-amino-2,3,4,6-tetradeoxy-D-glucose 4-*N*,*N*-dimethyltransferase

1. INTRODUCTION

Deoxyaminosugars comprise an important class of deoxysuagr moieties synthesized by a variety of organisms, including plants, fungi, and bacterial.[75] Deoxyaminosugars are usually formed via transamination reaction, which results in the substitution of a keto group with an amino group. The latter can be ionized under physiological pH, and therefore might be engaged in both electrostatic interactions with other ionizable groups and in the formation of hydrogen bonds with certain chemical groups on the target molecule. For example, certain macrolide antibiotics containing deoxyaminosugars bind to the peptidyl exit site of the ribosome and block the tunnel that channels nascent peptides away from the peptidel bond formation center.[40] For the 14-membered-ring macrolide erythromycin, it has been demonstrated that the deoxyaminosugar desosamine mediates hydrogen bonding interactions that are essential for binding of the antibiotic to the ribosome.[41] Alternatively, the amino group might be important for interactions between antibiotic molecules, as in the case of polyene macrolides. Such interactions appear to play an important role in the mechanism of action of this class of antibiotics, which is based on the ability of several antibiotic molecules to organize themselves into a channel spanning the cellular membrane of the target organism.[111] These features of deoxyaminosugars make studies of their synthesis

important in efforts to produce antibiotics with new deoxysugar appendages by reengineering biosynthetic pathways either *in vivo* or *in vitro*.

Several aminotransferases that catalyze the installation of an amino group on a sugar have been studied. TylB in the biosynthetic pathway of TDP-D-mycaminose, a sugar component of the antibiotic tylosin, produced by *Streptomyces fradiae*, has been shown to act as a C-3 aminotransferase by *in vitro* biochemical studies conducted in our laboratory.[129] DesI in the biosynthetic pathway of the D-desosamine portion of the antibiotics methymycin and neomethymycin (**87**) produced by *Streptomyces venezuelae*, is a C-4 aminotransferase catalyzing the conversion of TDP-6-deoxy-4-keto-glucose (**3**) to TDP-4-amino-4,6-dideoxy-glucose (**57**). The function of DesI has been assigned by both *in vitro* and *in vivo* studies (Scheme 4-1).[72] [135] DesV, also in the biosynthetic pathway of TDP-D-desosamine, has been shown to be a C-3 aminotransferase by



Scheme 4-1. Macrolides generated by wide type *Streptomyces venezuelae*, *des*I mutant and *des*V mutant.

successful isolation of the product, TDP-3-amino-3,4,6-trideoxy-glucose (**59**) from an enzymatic reaction of TDP-4,6-dideoxy-3-keto-glucose (**58**) and DesV.[189] Interestingly, the presence of both a C-3 aminotransferase and a C-4 aminotransferase in the same biosynthetic pathway clearly demonstrated the substrate regiospecificities of these enzymes. A *desI* knockout mutant of *Streptomyces venezuelae* failed to produce methymycin and neomethymycin (**87**). Instead, a methymycin analogue carrying D-quinovose was generated (**85**). The fact that *desV* did not complement the missing *desI* gene indicates that the C-3 aminotransferase, DesV, could not catalyze the C-4 transamination reaction, and thus, the biosynthesis of TDP-D-desosamine was stalled.[51,72] On the other hand, another *S. venezuelae* mutant in which *desV* gene was replaced by a thiostrepton resistance gene produced two macrolide compounds analogous to methymycin and neomethymycin (**87**), respectively, but both carrying a 4,6-dideoxy-glucose (**154**) instead of D-desosamine.[49] The fact that *desI* did not complement the missing *desV* gene means that the C-4 aminotransferase could not take a C-3 ketosugar as substrate, and the biosynthesis of D-desosamine was interrupted after the formation of the DesII product. A ketoreductase in the general cellular pool probably catalyzes the C-3 ketoreduction to form the C-3 hydroxyl group observed in the compounds isolated from the *desV* knockout mutant. It is likely that the substrate specificities demonstrated by DesI and DesV are general characteristics of aminotransferases. Sequence analysis revealed that all the assigned aminotransferases in the deoxysugar biosynthetic pathways could be divided into two groups (Figure 4-1): the C-3 aminotransferases of which DesV is an example, and C-4 aminotransferases, of which

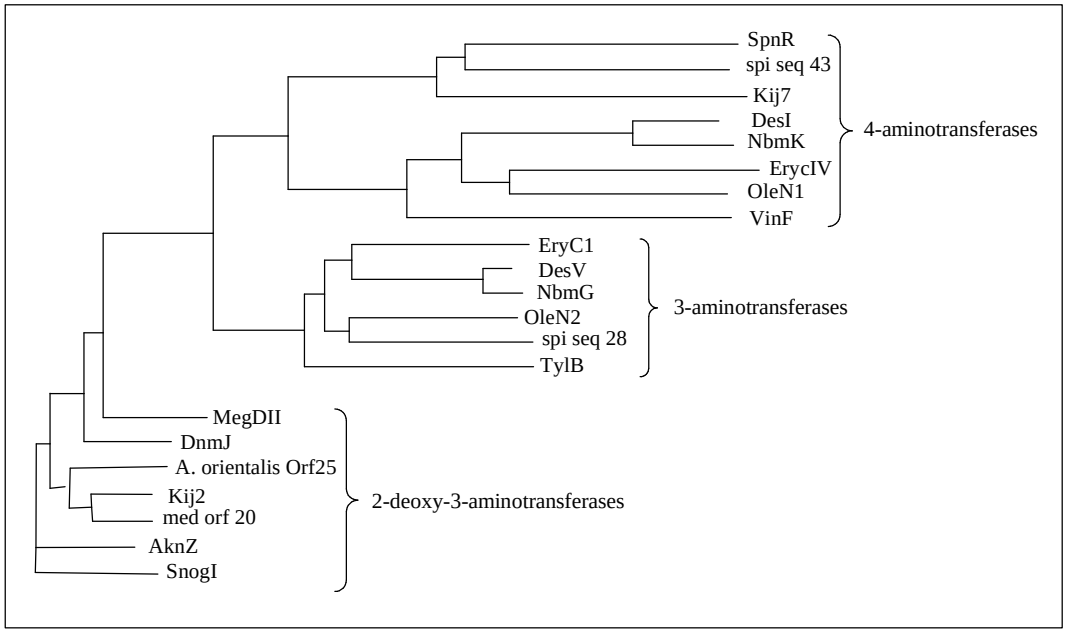
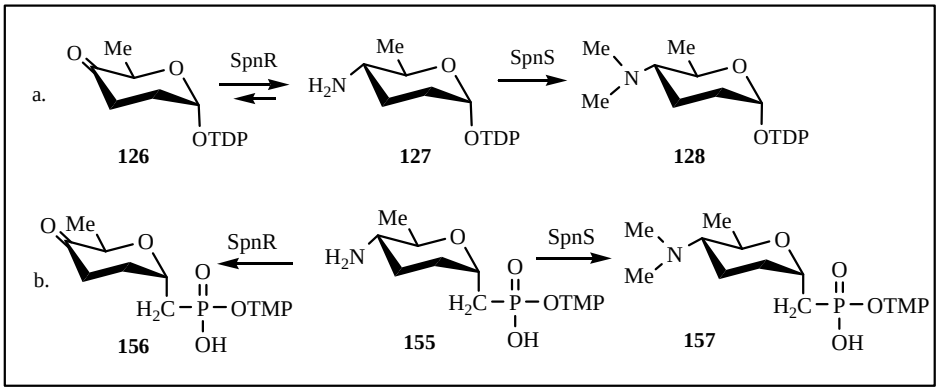


Figure 4-1 . Phylogenetic tree of aminotransferases involved in deoxysugar biosynthesis.

DesI is an example. The C-3 aminotransferases could be further divided into two subgroups: enzymes having 2-deoxy sugars as substrates and enzymes having 2-hydroxy sugars as substrates. The C-4 aminotransferases which are predicted to take 2-deoxy sugars as substrates and those that are predicted to take 2-hydroxy sugars do not fall into clear phylogenetic subclasses as do the C-3 aminotransferases. Within each group or subgroup, the sequences are very well conserved while the sequences of members of different groups have less similarity. The C-3 aminotransferases predicted to use 2-deoxy sugars as substrates are somewhat distantly related to other C-3 aminotransferases.

In the TDP-D-forosamine biosynthetic gene cluster, *spnR* is proposed to encode an aminotransferase catalyzing the conversion of TDP-2,3,6-trideoxy-4-keto-glucose (**126**) to TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**) (Scheme 4-2). However,



Schem 4-2. (a). SpnR and SpnS catalyzed reactions in the TDP-forosamine biosynthetic pathway. (b). SpnR and SpnS catalyzed reactions using an analogue of TDP-4-amino-2,3,4,6-tetradeoxy-glucose as substrate.

functional assignment solely based on sequence analysis can be misleading. For example, E_1 has significant sequence similarity to many aminotransferases, but its authentic function is to catalyze C-3 deoxygenation in the biosynthesis of CDP-ascarylose. SpnQ also shows considerable sequence similarity to a number of aminotransferases. However, conversion of TDP-2,6-dideoxy-4-keto-glucose (**24**) to TDP-4-amino-2,4,6-trideoxy-glucose (**153,** Figure 3-8) was shown to be only a secondary function of SpnQ. The C-4 transamination reaction catalyzed by SpnQ was kinetically less efficient than those catalyzed by typical aminotransferases. The major function of SpnQ was shown to be the C-3 deoxygenation of TDP-2,6-dideoxy-4-keto-glucose (**24**) in the biosynthetic pathway of TDP-D-forosamine. Thus, the least ambiguous and most

reliable way to assign an enzyme as an aminotransferase is by characterizing either gene knockout products or enzymatic reaction products.

Much information about C-3 aminotransferases has been obtained from the thorough biochemical studies conducted on TylB. In contrast, our knowledge on C-4 aminotransferases is still limited. Although the function of DesI has been confirmed by *in vivo* deletion experiments and *in vitro* enzymatic reactions, a detailed biochemical characterization was not carried out.

In this chapter, the aminotransferase function of SpnR is studied by *in vitro* enzymatic reactions, and a detailed biochemical characterization is described. The work done in Chapter 2 and Chapter 3 has clearly established the order of the C-2 and C-3 deoxygenation. Since both the C-2 and C-3 deoxygenations are activated by the presence of the C-4 keto group, it can be inferred that the C-4 transamination reaction should occur after the two deoxygenation reactions. The SpnR substrate, which is also the product of the SpnQ reaction, is TDP-2,3,6-trideoxy-4-keto-glucose (**126**). The compound has been isolated from a preparative SpnQ enzymatic reaction, but is unstable, and can easily decompose by losing TDP. In view of the instability of this compound and the reversibility of the transamination reaction, the SpnR catalyzed reaction was analyzed in the reverse direction using α-ketoglutarate and a chemically synthesized compound **155** as substrates (Scheme 4-2). Compound **155** is structurally similar to TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**) except that the oxygen bridging the anomeric carbon and the β-phosphor of TDP in TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**) is replaced with a methylene group.

Methylation reactions occur widely in nature. The substrates of methylations represent a broad range of biomaterials including DNA, tRNA, proteins, polysaccharides and small molecules. DNA methylation has been recognized for its prominent role in controlling diverse immune processes, tumor genesis, and age-related changes.[168,169] The cytosine residues in CG dinucleotides are the main sites for DNA methylation.[170,171] Studies have shown that DNA methylation precedes and directs histone modification to effect changes in chromatin structure. Histone methylation is chemically stable and may be reversible.[172,173] Histone methylation mainly occurs on lysine residues although arginine methylation has also been documented.[174,175] The specific function of histone methylation is largely decided by the location of the substrate residue.[171] Compared with the extensive studies on macromolecule methylation, methylations of macrolide antibiotics are less well studied although it has been documented that methylation of macrolide antibiotics plays an important role in macrolide antibiotic recognition specificities.[176]

Most methylation reactions are enzyme-catalyzed and SAM dependent. However, in some cases, the methyl donor can be coenzyme B_{12}, such as in methylation reactions catalyzed by one type of methionine synthases.[190,191] In this reaction, the methyl group of 5-methyltetrahydrofolate is passed to a reactive form of vitamin B_{12} to form methylcobalamin which can transfer the methyl group to the thiol side chain of homocysteine to form methionine. Macrolide antibiotic methylation can occur either on the sugar appendages or on the aglycone, and the methylation sites are usually nucleophilic NH_2-, (NHR-), OH-, and CHR_2-. Correspondingly, the methyltransferases

can be subclassified as *N*-methyltransferases, *O*-methyltransferases, and *C*-methyltransferases. Sequence analysis supports and illustrates these subclassifications. The proteins from different subclasses are only distantly related, while those within a given subclass are more closely related. *N*-methyltransferases can be further divided into two subgroups (Figure 4-2): 4-amino-*N*-methyltransferases of which SpnS from the TDP-D-forosamine biosynthetic way and VinG from the TDP-vicenisamine biosynthetic pathway are the representatives,[156] and 3-amino-*N*-methyltransferases, of which TylM1, DesVI, and several others, are examples.

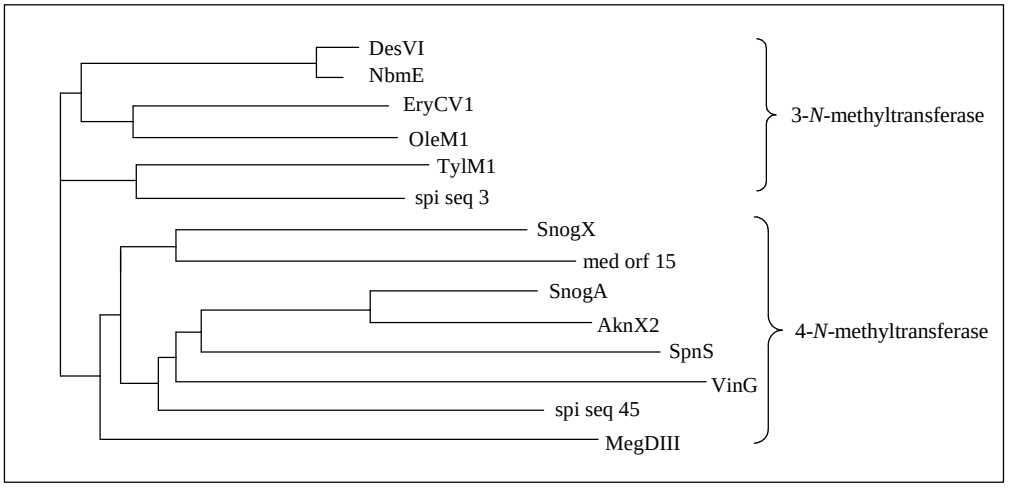


Figure 4- 2. Phylogenetic tree of 3-*N*-methyltransferases and 4-*N*-methyltransferases that are involved in deoxysugar biosynthesis.

Like other enzymes in the deoxysugar biosynthetic pathways, most methyltransferases are assigned based on sequence analysis since few of them have been biochemically characterized. TylM1, the 3-amino-*N*,*N*-dimethyltransferase in the biosynthetic pathway of TDP-D-mycaminose in antibiotic tylosin produced by

Streptomyces fradiae, has been studied in detail by our laboratory.[195] It has been demonstrated that the *N*,*N*-dimethylation is a stepwise nucleophilic process. The function of 3-*N*,*N*-dimethyltransferase DesVI has also been confirmed by enzymatic assays.[195] Among *C*-methyltransferases, TylC3 in the biosynthetic pathway of TDP-L-mycarose has been proved to be the enzyme catalyzing the conversion of TDP-2,6-dideoxy-4-keto-glucose (**28**, Scheme 1-2) to TDP-2,6-dideoxy-4-keto-3-*C*-methyl-glucose (**30**, Scheme 1-2).[167] The presence of the 4-keto group activates the α carbon and facilitates the proton abstraction on C-3. The resulting carbanion can serve as a good nucleophile in the *C*-methylation reaction. SAM has been shown to be the methyl donor in all of the characterized sugar methylation reactions.

In the biosynthetic pathway of TDP-D-forosamine, SpnS is suggested to catalyze 4-amino-*N*,*N*-dimethylation reaction (Scheme 4-2). Thus, the SpnS catalyzed reaction must occur later than the SpnR catalyzed transamination reaction. Since the SpnR reaction must be after the C-2 and C-3 deoxygenations, SpnS must be the last step in TDP-D-forosamine biosynthetic pathway. In this chapter, *in vitro* enzymatic activity studies and biochemical characterization on SpnS will be presented. The functional analysis of SpnR and SpnS described in this chapter has led to the complete establishment of the biosynthetic pathway of TDP-D-forosamine.

2. EXPERINETAL PROCEDURES

Materials. Most materials used in these experiments were identical to those described in Chapter 2. Other reagents, such as SAM, pyruvate, and α-ketoglutarate, were purchased from Sigma-Aldrich. All the chemicals used in the chemical synthesis of 155 were analytical grade or the highest quality commercially available.

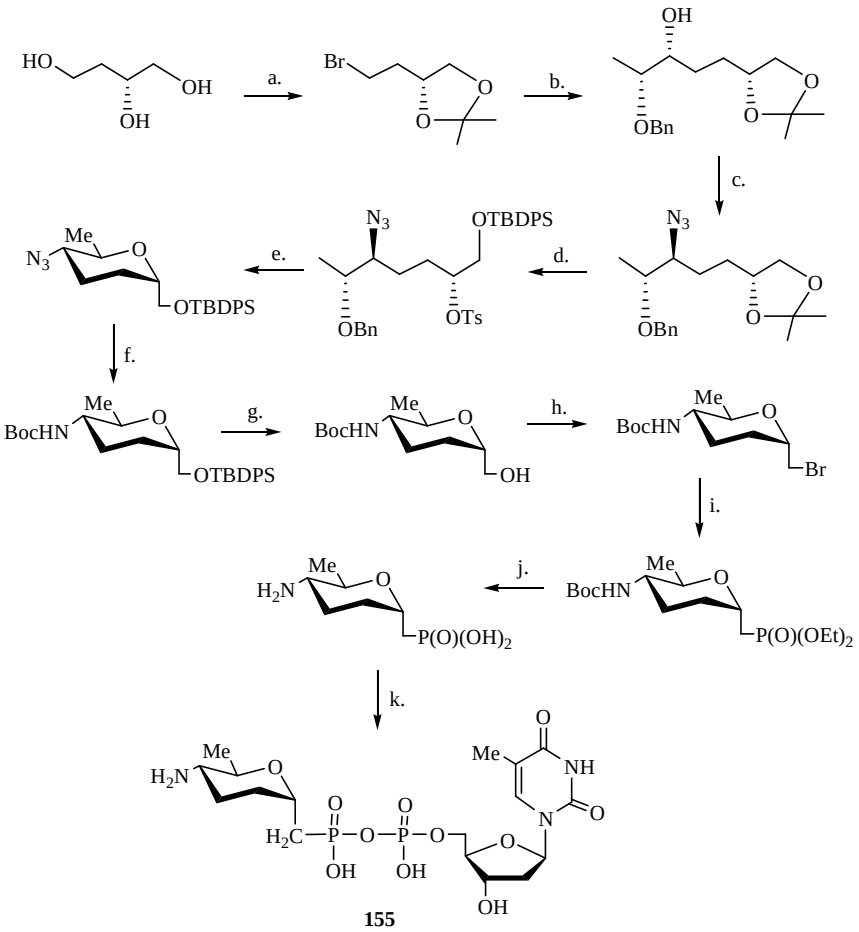
General Equipment. The equipment and other experimental apparati were identical to those described in Chapter 2.

Bacterial Strains. Various *E. coli* host cells used were identical to those described in Chapter 2.

Synthesis of 155. Chemical synthesis of 155, an analog of TDP-4-amino- 2,3,4,6-tetradeoxy-glucose (127) was performed by Dr. Kent Zhao following the scheme delineated in Scheme 4-3.

PCR Amplification of the spnR Gene. The cloning of *spnR* gene was completed by Dr. Kent Zhao. Two oligonucleotide primers complementary to the sequence at each end of the *spnR* gene were designed to amplify the gene for cloning into the pET24b(+) vector. The start primer (KZR-1), 5'-GGGGCCTCCATATGTGATCAACCTGCA-CCAG-3', contains an *Nde*I restriction site (in bold) and the first 18 nucleotides of the *spnR* gene. The halt primer (KZR-2) 5'-ATCCGCTCGAGCTTTCGGAGTGGTGA-TCTTTG-3', was designed to introduce an *Xho*I restriction site (in bold) immediately downstream of the coding sequence. The natural stop codon of *spnR* was not included in the sequence of KZR-2 in order to produce a *C*-terminal His_6-tagged fusion protein upon cloning into pET24b(+). Polymerase mediated amplification was carried out in a 0.5 mL microcentrifuge tube. To the reaction mixture was added 37.5 μL of double-deionized

H_2O, 5.0 μL of *pfu* polymerase buffer (10 x), 4.0 μL of deoxyribonucleotidyl-triphosphate mix (2.5 mM each), 1.0 μL of start primer KZR-1 (10 μM), 1.0 μL of halt primer KZR-2 (10 μM), 0.5 μL of genomic DNA (about 0.1 μg) from *Saccharopolyspora*



a) 1. Acetone, p-TsOH, 2. CBr_4, PPh_3, 55%; b). 1. Mg, THF, 2. (R)-2-benzyloxypropionaldehyde, 60%; c) DPPA, DEAD, PPh_3, 73%; d) 1. H^+ resin, 2. TBDPSCl, imidazole, 3. TsCl, 75%; e) 1. BCl_3, 2. NaH, DMSO, 52%; f) Pd/C, H_2, Boc_2O, EtOAc, 87%; g) TBAD, THF, 92%; h) CBr_4, PPh_3, CH_2Cl_2, 56%; i) $P(OEt)_3$, Bu_4NI, 85%; j) TMSBr, 89%; k) TMP-morpholidate, 1-*H*-tetrazole, pyridine, 30%.

Scheme 4-3. Synthetic scheme for compound 155, analog of TDP-4-amino-2,3,4,6-tetradeoxy-glucose (127).

spinosa as the amplification template, and 1.0 μL of cloned *pfu* polymerase (2.5 units). The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following thermal cycles: (1) 1 denaturation cycle of incubation at 95 °C for 3 min; (2) 30 amplification cycles of incubation at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 2.5 min; and (3) 1 extension cycle of incubation at 72 °C for 10 min. The tube was held at 4 °C prior to being removed from the thermal cycler.

PCR Amplification of the spnS Gene. The cloning of *spnS* gene was completed by Dr. Kent Zhao. Two oligonucleotide primers complementary to the sequence at each end of the *spnS* gene were designed to amplify the gene for cloning into the pET24b(+) vector. The start primer (KZS-1), 5'-GGGAATTCCATATGTCGCGCGTGAGCGA CAC-3', contains an *Nde*I restriction site (in bold) and the first 20 nucleotides of the *spnS* gene. The halt primer (KZS-2) 5'-ATCCGCTCGAGGTTGCGGATTCCGACGAAC-3', was designed to introduce an *Xho*I restriction site (in bold) immediately downstream of the coding sequence. The natural stop codon of *spnS* was not included in the sequence of KZS-2 in order to produce a *C*-terminal His_6-tagged fusion protein when cloned into pET24b(+). Polymerase mediated amplification was carried out in a 0.5 mL microcentrifuge tube. To the reaction mixture was added 37.5 μL of double-deionized H_2O, 5.0 μL of *pfu* polymerase buffer (10 x), 4.0 μL of deoxyribonucleotidyltriphosphate mix (2.5 mM each), 1.0 μL of start primer KZS-1 (10 μM), 1.0 μL of halt primer KZS-2 (10 μM), 0.5 μL of genomic DNA (about 0.1 μg) from *Saccharopolyspora spinosa* as the amplification template, and 1.0 μL of cloned *pfu* polymerase (2.5 units). The reaction mixture was then overlaid with a layer of mineral oil and subjected to the following

thermal cycles: (1) 1 denaturation cycle of incubation at 95 °C for 3 min; (2) 30 amplification cycles of incubation at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 2.5 min; and (3) 1 extension cycle of incubation at 72 °C for 10 min. The tube was held at 4 °C prior to being removed from the thermal cycler.

Cloning of the spnR and spnS genes. After PCR amplification, standard recombinant DNA techniques were used to clone the genes into an appropriate expression vector. Typically, 5 μL of the reaction mixture was withdrawn and electrophoresed on a 0.7% agarose gel to check the size and the yield of the product. The PCR product mixture was separated from the oil and subjected to DNA purification using QIAquick spin kit. After digestion with appropriate restriction enzymes, the entire mixture was loaded onto a 0.7% agarose gel for electrophoresis. The desired DNA band with expected size was excised and purified by the QIAquick spin kit. Ligation of the target gene with predigested pET24b(+) was achieved by the action of high concentration T4 DNA ligase. Typically, a ligation reaction of 20 μL including 5.0 unit of ligase was carried out for 2 h at room temperature.

Transformation of 20 μL *E.coli* DH5α host cells was carried out with 20 ng of the ligation product. The resulting cells were incubated on ice for 30 min followed by a heat shock at 42 °C for 90 s. The cells were then cooled down by incubation on ice for another 2 min before 800 μL of LB medium was added. After incubating at 37 °C with shaking (200 rpm) for 45 min, the culture was briefly centrifuged (13,000 rpm, 1 min). An aliquot of 700 μL of the supernatant was removed. The small cell pellet was resuspended in the remaining medium and then plated out on LB plates supplemented

with 30 μg/mL kanamycin. The plates were incubated for 20 h at 37 °C before six clonies were picked to inoculate 5 mL of liquid medium, which was incubated at 37 °C overnight. The plasmid DNAs were isolated from each of the cultures using QIAprep spin Miniprep kit. The plasmid DNA (40 ng) from each miniprep was digested with appropriate restriction enzymes. Positive clones were identified by the presence of the fragment with the expected size on the agarose gel after electrophoresis and were further confirmed by DNA sequencing. The plasmid DNA of the positive clones, pKZR and pKZS for *spnR* and *spnS*, respectively, were used to transform the *E. coli* BL21 (DE3) competent cells.

Purification of the SpnR Protein. All purification steps were carried out at 4 °C.

Step 1: Growth E. coli BL21 (DE3)-pKZR cells. An overnight culture of *E. coli* BL21 (DE3)-pKZR, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of three 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.3. The incubation temperature was adjusted to 24 °C and the incubation was continued for one more hour before the OD_{600} reached 0.4. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 15 h at 24 °C. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with 1 min

cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. An 8 mL of Ni-NTA slurry (4 mL of resin) was added to the crude extract, and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 30 mL of the lysis buffer followed by various amounts of wash buffers. The wash buffers and elution buffer used contained the same ingredients as the lysis buffer except that the concentrations of imidazole were altered. The column was washed sequentially with 30 mL of wash buffer containing 20 mM imidazole, 30 mL of wash buffer containing 40 mM imidazole, and 15 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffers passed through the column by gravity at a flow rate of 0.5 mL/min. After washing, the SpnR protein was eluted with 15 mL of elution buffer containing 100 mM imidazole. Eluted protein was concentrated to about 3 mL by ultrafiltration in an Amicon concentrator using a YM 10 membrane (Amicon).

Step 4: Sephacryl S-200 HR Chromatography. The protein from step 3 was applied to a Sephacryl S-200 HR column (2.5 x 100 cm) preequalibrated with buffer A (25 mM Tris·HCl, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The column was eluted with the same buffer at a flow rate of 0.8 mL/min. The fraction size of the eluent

was set at 8 mL. Pure fractions were identified by SDS-PAGE, and concentrated to 3.0 mL by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon). The purified SpnR protein was stored at –80 °C.

Purification of the SpnS protein. All purification steps were carried out at 4 °C.

Step 1: Growth E. coli BL21 (DE3)-pKZS cells. An overnight culture of *E. coli* BL21 (DE3)-pKZS, grown in LB medium supplemented with kanamycin (30 μg/mL) at 37 °C, was diluted 250-fold by addition of 4 mL of the overnight culture to each of six 1 L flasks of the same medium and incubated at 37 °C until the OD_{600} reached 0.4. The incubation temperature was adjusted to 30 °C and the incubation was continued for another half an hour before the OD_{600} of the culture reached 0.5. The culture was then induced with IPTG to a final concentration of 0.2 mM and allowed to grow for an additional 7 h at 30 °C. The cells were then harvested by centrifugation at 5,000 g for 10 min at 4 °C and stored at –80 °C. The typical yield of wet cells was about 6 g per liter.

Step 2: Crude Extract Preparation. The cells were resuspended in 80 mL of lysis buffer (50 mM sodium phosphate, 300 mM NaCl, 5 mM imidazole, 15% glycerol, 5 mM 2-mercapto-ethanol, pH 8.0), and disrupted with five 1 min sonication bursts, with 1 min cooling intervals between bursts. Cellular debris was removed by centrifugation at 35,000 g for 20 min. The supernatant was transferred into a large centrifuge bottle for incubation with Ni-NTA resin.

Step 3: Ni-NTA Affinity Chromatography. To the crude extract, was added 8 mL of Ni-NTA slurry (4 mL of resin) and the resulting mixture was incubated for 2 h with gentle shaking at 4 °C. The majority of the supernatant was decanted after the resin was

spun down to the bottom of the centrifuge bottle with brief centrifugation at low speed (1,000 g). The remaining resin slurry was mixed well and loaded into an empty glass column (1 x 25 cm). The column was washed with 30 mL of lysis buffer followed by various amounts of wash buffers. The wash buffers and elution buffer used contained the same ingredients as the lysis buffer except that the concentrations of imidazole were altered. The column was washed sequentially with 30 mL of wash buffer containing 20 mM imidazole, 30 mL of wash buffer containing 40 mM imidazole, and 15 mL of wash buffer containing 60 mM imidazole. No peristaltic pump was used. The buffer passed through the column by gravity at a flow rate of 0.5 mL/min. After washing, the SpnS protein was eluted with 15 mL of elution buffer containing 100 mM imidazole. Eluted protein was concentrated to about 3 mL by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon).

Step 4: Sephacryl S-200 HR Chromatography. The protein from the previous step was applied to a Sephacryl S-200 HR column (2.5 x 100 cm) preequalibrated with buffer A (25 mM Tris·HCl, 1 mM 2-mercapto-ethanol, 15% glycerol, pH 7.5). The column was eluted with the same buffer at a flow rate of 0.8 mL/min. The fraction size was set at 8 mL. Pure fractions were identified by SDS-PAGE, and concentrated to 3.0 mL by ultrafiltration in an Amicon concentrator using a YM-10 membrane (Amicon). The purified SpnS protein was stored at –80 °C.

Polyacrylamide Gel Electrophoresis. The subunit molecular mass and the purity of the protein samples were assessed by SDS-PAGE as described in Chapter 2.

Molecular Weight Determination. The native molecular weights of the purified protein samples were determined by gel filtration chromatography following the procedure described in Chapter 2.

Activity Assay for SpnR. An HPLC-based assay was developed for SpnR with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) connected to a CarboPac™ PA1 guard column (4 x 50 mm) by monitoring the consumption of the substrate and the formation of a new compound at 267 nm. A typical assay mixture of 80 μL contained 0.5 mM of the substrate **155**, 30 mM of α-ketoglutarate, 2.0 mM of DTT, 2.0 mM of $MgCl_2$, 0.2 mM of PLP, 10 μM of SpnR, and 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of SpnR and was carried out at room temperature (24 °C). A suitable amount of reaction mixture was withdrawn at appropriate time intervals. The reaction was terminated by dilution with a 10-fold excess of water and flash freezing in liquid nitrogen. The enzyme was removed by filtering through a Centricon YM-10 microconcentrator by centrifugation. An appropriate amount of the filtrate was subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60% B over 20 min was used to resolve the substrate **155** (retention time = 6.5 min) and the SpnR product (retention time = 31.3 min). Buffer A was water and Buffer B was 500 mM NH_4OAc, pH 7.0, adjusted with diluted $NH_3 \cdot H_2O$. The elution gradient was followed by a final wash with 100% B for 5 min and reequilibration with 5% B for 15 min. The absolute rate of SpnR product formation was calculated as the number of nmols of product formation per min, and percent conversion of the substrate was used directly to determine the relative rate.

Reconstitution of SpnR with PLP under Native Conditions. A 200 μL solution containing 120 μM SpnR, 2.0 mM PLP, 1 mM DTT, and 50 mM potassium phosphate (pH 7.5) was incubated at room temperature for 6 h. The entire solution was then loaded onto an Econo-Pac 10 DG column (Bio-Rad). The reconstituted protein was eluted with 50 mM potassium phosphate buffer (pH 7.5). Fractions containing the SpnR protein were pooled and concentrated. Any residual unbound PLP was removed by repeated dilution and concentration using Centricon YM-10 microconcentrator. The presence of the enzyme bound PLP was determined by UV-visible spectroscopy.

Reconstitution of SpnR with all Necessary Substrates Present. A 200 μL solution containing 120 μM SpnR, 2.0 mM PLP, 1 mM DTT, 0.5 mM of the substrate 155, 5.0 mM α-ketoglutarate, and 50 mM potassium phosphate buffer (pH 7.5) was incubated at room temperature for 6 h. The entire solution was then loaded onto Econo-Pac 10 DG column (Bio-Rad). The reconstituted protein was eluted with 50 mM potassium phosphate buffer (pH 7.5). Fractions containing the SpnR protein were pooled and concentrated. Any residual unbound PLP was removed by repeated dilution and concentration using Centricon YM-10 microconcentrator. The presence of enzyme bound PLP was determined by UV-visible spectroscopy.

UV-visible Spectrum of the SpnR Protein. The absorbance of the as-isolated SpnR protein and the reconstituted SpnR protein (reconstituted either in the presence or in the absence of the substrates) was measured from 250 nm to 750 nm on a Beckman DU-650 spectrophotometer.

Effect of Reconstitution on SpnR Catalytic Activity. PLP was omitted from the SpnR activity assay mixture while the concentrations of other reagents were held constant and the as-isolated SpnR was replaced with the reconstituted SpnR (reconstituted either in the presence or in the absence of the substrates). The substrate conversion was compared with that of the reaction catalyzed by the as-isolated SpnR.

Mg^{2+} Requirement in SpnR Catalysis. The Mg^{2+} ion was omitted from the SpnR activity assay mixture and the relative reaction rate was compared with that of a reaction containing Mg^{2+}.

PLP Concentration Effect. While the concentrations of other reagents were held constant, the concentration of PLP in the SpnR activity assay mixture was set at either 0, 10 μM, 20 μM, 40 μM, or 80 μM. The relative reaction rates were compared to determine the PLP concentration effect.

PMP as the Cofactor. PLP was replaced with PMP in the SpnR activity assay. The competence of PMP as cofactor in the transamination reaction was evaluated with that of PLP by comparison of the substrate conversions of each reaction.

Determination of the Optimal Amino Acceptor. Two α-keto acids, pyruvate and α-ketoglutarate were examined for their efficiency as the amino acceptor in the SpnR catalyzed reaction. The α-keto acids were added to a final concentration of 3.0 mM in assays described previously. The efficiency of these two α-keto acids as amino acceptors was examined by comparing the relative reaction rates when using each in the reaction.

Optimal pH for SpnR Catalysis. The optimal pH for the reaction catalyzed by SpnR was examined in buffers with different pHs. The concentrations of the reaction

reagents were held constant, while the pH was varied over a range of 5.5 to 9.5. A 50 mM MES buffer was used in the reaction at pH 5.5; a 50 mM Bis-Tris-Propane buffer was used in the reaction at pH 6.5; a 50 mM potassium phosphate buffer was used in the reaction at pH 7.5; and a 50 mM Bis-Tris-propane buffer was used for the reactions at pH 8.5 and 9.5. The relative reaction rates at different pHs were compared to determine the optimal pH of the SpnR reaction.

Isolation and Characterization of the SpnR Product. The reaction was carried out in the reverse direction, from aminosugar to ketosugar, and the substrate used was 155, an analog of TDP-4-amino-2,3,4,6-tetradeoxy-glucose. A preparative scale incubation was carried out in which 7.8 mM 155, 143 mM α-ketoglutarate, 0.3 mM PLP, and 0.23 mM SpnR were incubated in 3 mL of 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of SpnR and was carried out at 24 °C. The progress of the incubation was monitored by HPLC as described previously for the SpnR activity assay. The reaction was almost complete (97%, based on peak area integration) after 15 h. The enzyme was removed using a Centricon YM-10 and the filtrate was applied onto a gel filtration column (Bio-Rad P2, extra fine, 2.5 x 150 cm) preequalibrated with 25 mM NH_4HCO_3. The compounds were eluted with the same buffer at a flow rate of 12 mL/h at 4 °C. Fractions were assayed spectrophotometrically by scanning from 250 nm to 400 nm. The fractions that displayed a single absorption maximum at 267 nm were lyophilized as individual fractions and analyzed by NMR. Spectra data for the SpnR reverse reaction product: ^{1}H NMR (D_2O, 4-keto and 4-keto hydrated form): δ 1.11 (d, J = 4.2 Hz, 5-Me), 1.21 (d, J = 4.2 Hz, 5-Me), 1.84 (3H, m, 5’’-Me), 1.57-2.58 (6H, m, 2-H,

3-H, CH_2-P), 2.28 (2H, m, 2'-H), 3.64 (1H, q, *J* = 3.9 Hz, 5-H), 4.20 (1H, q, *J* = 3.9 Hz, 5-H), 4.03-4.35 (4H, m, 4'-H, 5'-H, 1-H), 4.52 (1H, ddd, 3'-H), 6.26 (1H, t, *J* = 4.2 Hz, 1'-H), 7.66 (s, 6''-H), 7.67 (s, 6''-H). ^{13}C NMR (D_2O, hydrated form): 12.0, 15.2, 27.8, 29.2, 33.6 (d, *J* = 138.2), 38.8, 65.2, 67.6, 71.4, 75.2, 85.2, 85.6, 92.8, 112.0, 139.8, 152.0, 167.5. ^{31}P NMR (D_2O, hydrated form): 14.8 (d, *J* = 26 Hz), 13.9, (d, *J* = 26 Hz), -10.6 (d, *J* = 26 Hz). High resolution FAB-MS calculated for $C_{17}H_{26}N_2O_{12}P_2$ $(M-H)^-$ was 511.0883, and a mass of 511.0880 was found.

Assay of SpnR in the Forward Direction. After the SpnR reverse reaction product was isolated, it became possible to assay SpnR in the forward direction. The substrate is an analog (156) of TDP-2,3,6-trideoxy-4-keto-glucose (126). An assay mixture of 50 μL contained 0.6 mM 156, 12.0 mM L-glutamate, 1.0 mM $MgCl_2$, 1.0 mM DTT, 50.0 μM PLP and 10 μM SpnR. After the reaction was incubated at 24 °C for 3 h, the enzyme was removed using a Centricon YM-10 microconcentrator and the filtrate was subjected to HPLC analysis. The elution program was the same as that used for the SpnR activity assays.

Assay of SpnR in the Forward Direction with Authentic Substrate. The highly deoxygenated and unstable SpnQ product (126) was generated immediately prior to use. A typical assay mixture of 80 μL contained 0.70 mM SpnN product (24), 0.15 mM PMP, 30 μM ferredoxin, 30 μM ferredoxin reductase, 2.0 mM NADPH, 30 μM SpnQ, and 50 mM potassium phosphate buffer (pH 7.5). After the incubation was carried out at room temperature for 3 h, all the enzymes in the reaction mixture were removed using a Centricon YM-10 concentrator. To the filtrate (80 μL) were added 0.025 μmol of PLP,

1.0 μmol of L-glutamate and 0.1 μmol of $MgCl_2$. The reaction mixture was incubated at room temperature for 2 h. The enzyme SpnR was then removed using a Centricon YM-10 microconcentrator and the final reaction filtrate was subjected to HPLC analysis. The separation program was the same as the one used for the SpnR activity assays.

Substrate Specificity of SpnR. The SpnN product (24) was examined as a possible substrate for SpnR. A typical assay mixture of 25 μL contained 1.05 mM SpnN product (24), 0.25 mM PLP, 10.0 mM L-glutamate, 30 μM SpnR and 50 mM potassium phosphate buffer (pH 7.5). After the assay mixture was incubated at 37 °C for 18 h, the enzyme was removed by filtering through a Centricon YM-10 microconcentrator by centrifugation and the filtrate was subjected to HPLC analysis. In another incubation, 26.2 nmol of the SpnN product was incubated with 12.5 nmol of PLP, 15 nmol of L-glutamate and 0.74 nmol of SpnQ in 25 μL of 50 mM potassium phosphate buffer (pH 7.5). SpnQ, under these conditions, was previously shown to act as a 4-aminotransferase, as described in Chapter 3. After the reaction mixture was incubated at 37 °C for 10 h, the enzyme was removed using a Centricon YM-10 microconcentrator. Under these conditions, the SpnQ-catalyzed 4-aminotransfer did not proceed to completion, and thus, there was still some SpnN product present in the reaction mixture after removal of SpnQ. To the filtrate were added 7.5 nmol of PLP, 0.25 μmol of L-glutamate and 1.5 nmol of SpnR and the incubation was carried out at either room temperature or 37 °C for 4 h. After the enzyme was removed using a Centricon YM-10 microconcentrator, the filtrate was analyzed by HPLC. The elution program was the same as the one used for the SpnR activity assays. This incubation with SpnR was performed to test whether SpnQ and

SpnR were both catalyzing the identical reaction under the conditions used. If so, this would provide further evidence supporting the identification of the product as TDP-4-amino-2,4,6-trideoxy-glucose (153).

Determination of the Equilibrium Constant of the SpnR Reaction. A 50 μL of reaction mixture contained 0.6 mM 155, 0.6 mM α-ketoglutarate, 50 μM PLP, 10 μM SpnR and 50 mM potassium phosphate buffer (pH 7.5). After the reaction was carried out at room temperature for 12 h, the enzyme was removed using Centricon YM-10 microconcentrator, and the filtrates were subjected to HPLC analysis. The elution program was the same as the one used for the SpnR activity assays. The concentrations of the keto- and aminosugar were calculated based on the integrals of the corresponding peaks. The equilibrium constant K_{eq} was defined as in Equation 4-1, where $[\alpha\text{-KG}]_0$ is the initial concentration of α-KG.

Activity Assay for SpnS. An HPLC-based assay was developed for SpnS with a Dionex CarboPac™ PA1 analytical column (4 x 250 mm) connected to a CarboPac™ PA1 guard column (4 x 50 mm) by monitoring the consumption of the substrates and the

$$Keq = \frac{[\text{Aminosugar}]\,[\text{a-KG}]}{[\text{Ketosugar}]\,[\text{Glu}]} = \frac{[\text{Aminosugar}]\,\{[\text{a-KG}]_0 - [\text{ketosugar}]\}}{[\text{Ketosugar}]^2}$$

Equation 4-1.

formation of new compounds at 267 nm. A typical assay mixture of 60 μL contained 1.05 mM **155**, 2.0 mM SAM, 2.0 mM $MgCl_2$, 2.0 mM DTT, 10 μM SpnS, and 50 mM potassium phosphate (pH 7.5). The reaction was initiated by the addition of SpnS and

was carried out at 37 °C. A suitable amount of the reaction mixture was withdrawn at appropriate time intervals and the reaction was terminated by denaturing the protein at 90 °C for 4 min. The enzyme was removed by filtering through a Centricon YM-10 microconcentrator with centrifugation and the filtrate was subjected to HPLC analysis. A linear gradient from 5 to 20% B over 15 min followed by another linear gradient from 20 to 60% B over 20 min while Buffer A is H_2O and Buffer B is 500 mM NH_4OAc (pH 7.0, adjusted with diluted $NH_3 \cdot H_2O$), was able to separate **155** (retention time = 6.5 min) and by-product *S*-adenosylhomocysteine (retention time = 3.0 min). However, SAM degradation product and the SpnS product overlapped with each other and could not be resolved. The elution gradient was followed by a final wash with 100% B for 5 min and a reequilibration with 5% B for 15 min.

Preparation and Characterization of the SpnS Product. A preparative scale incubation was carried out in which 16.5 μmol of **155**, 41.2 μmol of SAM, 4 μmol of $MgCl_2$, 4 μmol of DTT and 0.16 μmol of SpnS were incubated in 2 ml of 50 mM potassium phosphate buffer (pH 7.5). The reaction was initiated by the addition of SpnS and was carried out at 37 °C. The progress of the reaction was monitored by the HPLC assay described previously. After 15 h, about 50% of the substrate was converted to the methylation product as judged by peak area integration. The enzyme was then removed by filtering through a Centricon YM-10 membrane by centrifugation and the filtrate was applied to a gel filtration column (Bio-Rad P2, extra fine, 2.5 x 120 cm) preequalibrated with 25 mM NH_4HCO_3. The column was eluted with the same buffer at a flow rate of 12 mL/h at 4 °C. Fractions were assayed spectrophotometrically by scanning from 250 nm

to 400 nm. The fractions that displayed a single absorption maximum at 267 nm were lyophilized as individual fractions and analyzed by NMR. Spectra data for the SpnS reaction product: ^{1}H NMR (D_2O): δ 1.32 (3H, d, *J* = 4.2 Hz, 5-Me), 1.85 (3H, s, 5''-Me), 1.44-2.16 (6H, m, 2-H, 3-H, CH_2-P), 2.24-2.34 (2H, m, 2'-H), 2.66 (3H, s, NH-Me), 3.02 (1H, q, *J* =1.5 Hz, 4-H), 4.05-4.16 (5H, m, 4'-H, 5'-H, 1-H, 5-H), 4.54 (1H, ddd, 3'-H), 6.27 (1H, t, *J* = 3.9 Hz, 1'-H), 7.67 (1H, s, 6''-H). ^{13}C NMR (D_2O): δ 11.9, 15.4, 19.2, 26.0, 30.9, 34.2 (d, *J* = 138.2), 38.8, 57.7, 65.4, 66.1, 68.9, 71.1, 85.2, 85.6, 112.2, 138.0, 152.0, 167.2. ^{31}P NMR (D_2O): δ 14.1 (d, *J* = 27 Hz), -10.7 (d, *J* = 27 Hz). High-resolution FAB-MS calculated for $C_{18}H_{31}N_3O_{11}P_2$ $(M+H)^+$ was 528.1212, and a mass of 528.1502 was found.

Effects of Metals on SpnS Catalysis. Solutions of 2.0 mM of a variety of metals ions were tested for their effects on SpnS catalysis by including the appropriate salt of each individual ion Al^{3+}, Fe^{3+}, Zn^{2+}, Mg^{2+}, Fe^{2+}, Co^{2+}, and Ca^{2+} in the SpnS activity assay while maintaining the concentrations of other reagents unchanged. The elution program was the same as the one used for the SpnR activity assays.

Effect of Mg^{2+} Concentration on SpnS Catalysis. Solutions of 100 μM, 200 μM, 400 μM, 800 μM, 2.0 mM, and 10 mM of $MgCl_2$ were individually included in the SpnS activity assays. The relative reaction rates were compared by HPLC analysis. The elution program was the same as the one used for the SpnR activity assays.

Optimal pH for SpnS Catalysis. The optimal pH for the reaction catalyzed by SpnR was examined in buffers with different pHs. The concentrations of the reaction reagents were held constant, while the pH was varied over a range of 5.5 to 9.5. A 50

mM MES buffer was used in the reaction at pH 5.5; a 50 mM Bis-Tris-Propane buffer was used in the reaction at pH 6.5; a 50 mM potassium phosphate buffer was used in the reaction at pH 7.5; and a 50 mM Bis-Tris-propane buffer was used for the reactions at pH 8.5 and 9.5. The relative reaction rates at different pHs were compared to determine the optimal pH for the SpnS catalysis.

The Time Courses of Methylation. The time courses of mono-and di-methylation of **155** were determined by HPLC following the consumption of substrate and formation of monomethylated and dimethylated products. A typical assay mixture consisted of 1.0 mM **155**, 2.0 mM SAM, 2.0 mM $MgCl_2$, 2.0 mM DTT and 10.0 μM SpnS in 50 μL of 50 mM potassium phosphate buffer (pH 7.5). Aliquots were withdrawn at appropriate time intervals, diluted 20 fold, filtered through Centricon YM-10 by centrifugation to remove proteins, and injected to an Adsorbosphere SAX column (5μ, 4.6 x 250 mm) monitored at 267 nm. A linear gradient from 0 to 20% B between buffer A (50 mM potassium phosphate, pH 3.5) and buffer B (500 mM potassium phosphate, pH 3.5) over 20 min gave base line separation between **155** (retention time = 11.2 min) and the monomethylated product (retention time = 20.2 min). The peak of the dimethylated product still overlapped with those of SAM and its degradation products and therefore could not be directly observed. The elution gradient was followed by a final wash with 100% B for 5 min, and a reequilibration with 0%B for 15 min. The amounts of **155** and the monomethylated product were calculated from the integrals of the corresponding peaks. The amount of the dimethylated product was calculated by subtracting the

integrals of both the remaining substrate and the monomethylated product from that of the initial substrate.

Dimethylation Assay for SpnS. The monomethylated product (**158**) of SpnS isolated from the preparative scale reaction was used as the substrate to examine the dimethylation reaction catalyzed by SpnS. A typical assay mixture of 50 μL contained 1.0 mM of the monomethylation product (**158**), 2.0 mM SAM, 2.0 mM $MgCl_2$, 2.0 mM DTT, 20 μM SpnS and 50 mM of potassium phosphate (pH 7.5). The reaction was initiated by the addition of SpnS and was carried out at 37 °C for 12 h. The enzyme was removed by filtering through a Centricon YM-10 microconcentrator by centrifugation and the filtrate was subject to HPLC analysis (Dionex column). The elution program was the same as the one used for SpnR activity assay. As has been mentioned before, neither the monomethylation product nor the dimethylation product, could not be separated from SAM degradation products under the condition used. Thus the whole overlapped peaks were collected. The compounds were dried by lyophilization and redissolved in 100 μL of H_2O. The solution was submitted to MS analysis.

3. RESULTS AND DISCUSSION

Synthesis of an Analog of TDP-4-amino-2,3,4,6-tetradeoxy-glucose. The chemical synthesis of compound **155** was completed by Dr. Kent Zhao. In **155**, the oxygen atom that connects the hexose ring and TDP in the case of TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**) is replaced with a methylene group. Compound **155** was useful in both the SpnR and SpnS enzymatic reactions, serving as the substrate of both the

methylation reaction and the reverse transamination reaction. The compound was stable even being heated at 90 °C. The overall chemical synthesis took 11 steps. Spectra data for **155**: ^{1}H NMR (D_2O): δ 1.20 (3H, d, *J* = 6.9 Hz, 5-Me), 1.79 (3H, s, 5''-Me), 1.50-2.05 (4H, m, 2-H, 3-H), 1.90-2.02 (2H, m, CH_2-P), 2.12 (2H, m, 2'-H), 3.44 (1H, m, 4-H), 3.85 (1H, m, 5-H), 4.02 (4H, m, 4'-H, 5'-H, 1-H), 4.48 (1H, ddd, 3'-H), 6.21 (1H, t, *J* = 6.6 Hz, 1'-H), 7.67 (1H, s, 6''-H). ^{31}P NMR (D_2O): δ 14.3 (d, *J* = 27 Hz), -10.6 (d, *J* = 27 Hz). High-resolution FAB-MS calculated for $C_{17}H_{28}N_3O_{11}P_2$ $(M-H)^-$ was 512.1199, and a mass of 512.1189 was found.

Cloning and Expression of the SpnR and SpnS Proteins. In the biosynthetic gene cluster of TDP-D-forosamine, functions of *spnO*, *spnN* and *spnQ* encoded enzymes have been characterized in detail. The *spnR* and *spnS* genes are physically adjacent to these three genes. The proteins encoded by *spnR* and *spnS* were proposed based on sequence analysis to be an aminotransferase and a methyltransferase, respectively, catalyzing the last two steps in the biosynthetic pathway of TDP-D-forosamine. We decided to clone *spnR* and *spnS* to complete the characterization of the TDP-D-forosamine biosynthetic pathway. Cloning of *spnR* into the pET24b(+) vector gave a construct that was highly efficient in producing soluble SpnR, accounting for more than 20% of the total proteins in the crude extract. The *spnS* gene could also be well expressed when cloned into pET24b(+) vector, although a portion of the expressed proteins formed inclusion bodies. As judged by SDS-PAGE of the crude extracts of the host cells, SpnS was not a major soluble protein.

Purification and Characterization of the SpnR Protein (Figure 4-3). The protein was purified to homogeneity after Ni-NTA and Sephacryl S-200 chromatography. The typical yield was about 80 mg of pure SpnR from a 3 L culture. The protein had a light yellow color characteristic of PLP-containing enzymes. Its subunit molecular weight of 43 kDa as revealed by SDS-PAGE correlates well with the predicted value of 43358 Da

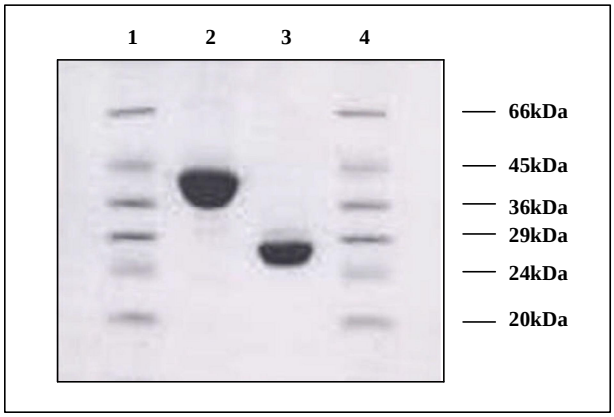


Figure 4-3. SDS-polyacrylamide gel electrophoresis of proteins obtained in Chapter 4. Lanes 1 and 4, molecular markers; lane 2, *C*-terminal His-tagged SpnR protein; lane 3, *C*-terminal His-tagged SpnS.

calculated from the deduced amino acid sequence (including the *C*-terminal His_6-tag). A M*r* of 85.9 kDa, as judged by gel filtration chromatography, indicates that the native SpnR exists as a homodimer. The concentrated SpnR protein is stable and can endure repeated freeze/thaw cycles without losing noticeable activity.

Effect of Mg^{2+} on the SpnR Catalysis. Mg^{2+} is a divalent cation commonly used in the reactions in which the substrate and/or product have a diphosphate functional group. The cation is supposed to stabilize the negatively charged phosphate groups and

thus facilitates the reaction. Therefore, Mg^{2+} was initially included in the SpnR activity assays. However, it was found that the relative reaction rate was not affected when Mg^{2+} was omitted from the reaction mixture. Thus, Mg^{2+} was not required in the SpnR reaction.

Effect of Different Concentrations of Exogenous PLP. PLP is the commonly used coenzyme in the transamination reactions. When different amounts of PLP were included in SpnR assay reactions, it was found that the relative reaction rate was greatly increased by the presence of PLP in a concentration-dependant manner. As shown in Figure 4-4, the reaction rate increased when the concentration of PLP was raised from 0 μM to 40 μM. Beyond this point, additional exogenous PLP could not further improve the reaction rate.

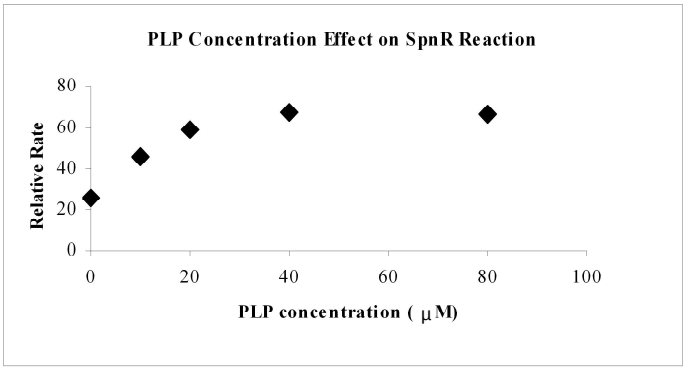


Figure 4-4. Effect of PLP concentration on SpnR reaction with 155 as the substrate.

Effect of Exogenous PMP. PMP is another commonly used coenzyme of aminotransferases. When PMP was included in the SpnR assay mixture, PMP demonstrated its capability to increase the reaction rate. The same amount of **155** was

converted to its 4-keto product as that in the reaction using PLP after the reaction reached equilibrium.

Reconstitution of SpnR with PLP under Native Conditions. The reconstituted protein had a stronger yellow color compared with the as-isolated SpnR indicating the presence of an increased amount of enzyme-bound cofactor.

Reconstitution of SpnR with all Necessary Substrates Present. In this reconstitution process, SpnR was incubated with PLP in the presence of **155** and α-ketoglutarate. The reconstituted protein showed a stronger yellow color compared with the as-isolated SpnR indicating the presence of an increased amount of enzyme-bound cofactor.

UV Scan of the as-Isolated SpnR and Reconstituted SpnR. The as-isolated protein showed little absorbance at 410 nm. The protein was estimated to contain 0.07 PLP per monomer calculated from the extinction coefficient of PLP ε_{410} = 6000 $M^{-1}cm^{-1}$.[192-194] In the UV-visible spectrum (Figure 4-5), the absorbance at 410 nm of the reconstituted enzymes, which were reconstituted either in the presence or absence of the substrates, was significantly enhanced compared with that of the as-isolated enzyme. The two methods of reconstitution did not show a difference at the level of PLP incorporation in reconstituted SpnR as judged by similar A_{410}/A_{280} values (0.12). The PLP content, 0.85 per monomer, is the same for the reconstituted enzymes obtained either way. Such results are different from those obtained from TylB biochemical studies. TylB is the

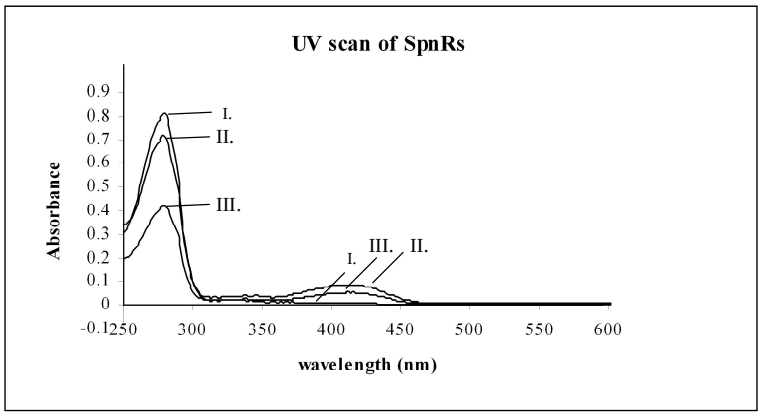


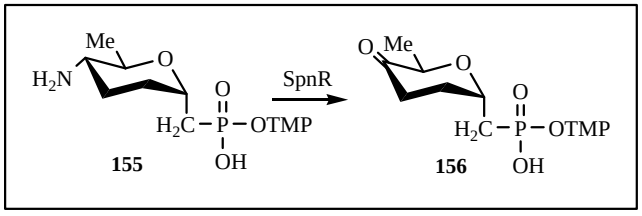
Figure 4-5. UV scan spectra of SpnRs. (I). the as-isolated SpnR. (II). SpnR reconsituted with PLP in the presence of 155 and -ketoglutarate. (III). SpnR reconstituted with PLP in the absence of 155 and -ketoglutarate.

aminotransferase in the biosynthetic pathway of TDP-mycaminose. The as-isolated TylB showed no absorbance above 300 nm in its UV-vis spectrum and was evidently devoid of any PLP coenzyme. The enzyme was not active in amino transfer reactions. Reconstitution of TylB in the absence of its substrates was not successful whether it was under the native condition or denaturing condition since the reconstituted enzyme still did not show any enzymatic activity. The active reconstituted enzyme was obtained in the presence of all the necessary substrates. Thus, SpnR was different from TylB in its coenzyme binding ability. This difference can be rationalized by the fact that TylB is a C-3 aminotransferase, while SpnR is a C-4 aminotransferase. The sequences of these two enzymes only have 26.9 % identity.

Effect of SpnR Reconstitution on its Catalytic Activity. It has been demonstrated that the as-isolated SpnR was capable of catalyzing the reverse transamination reaction

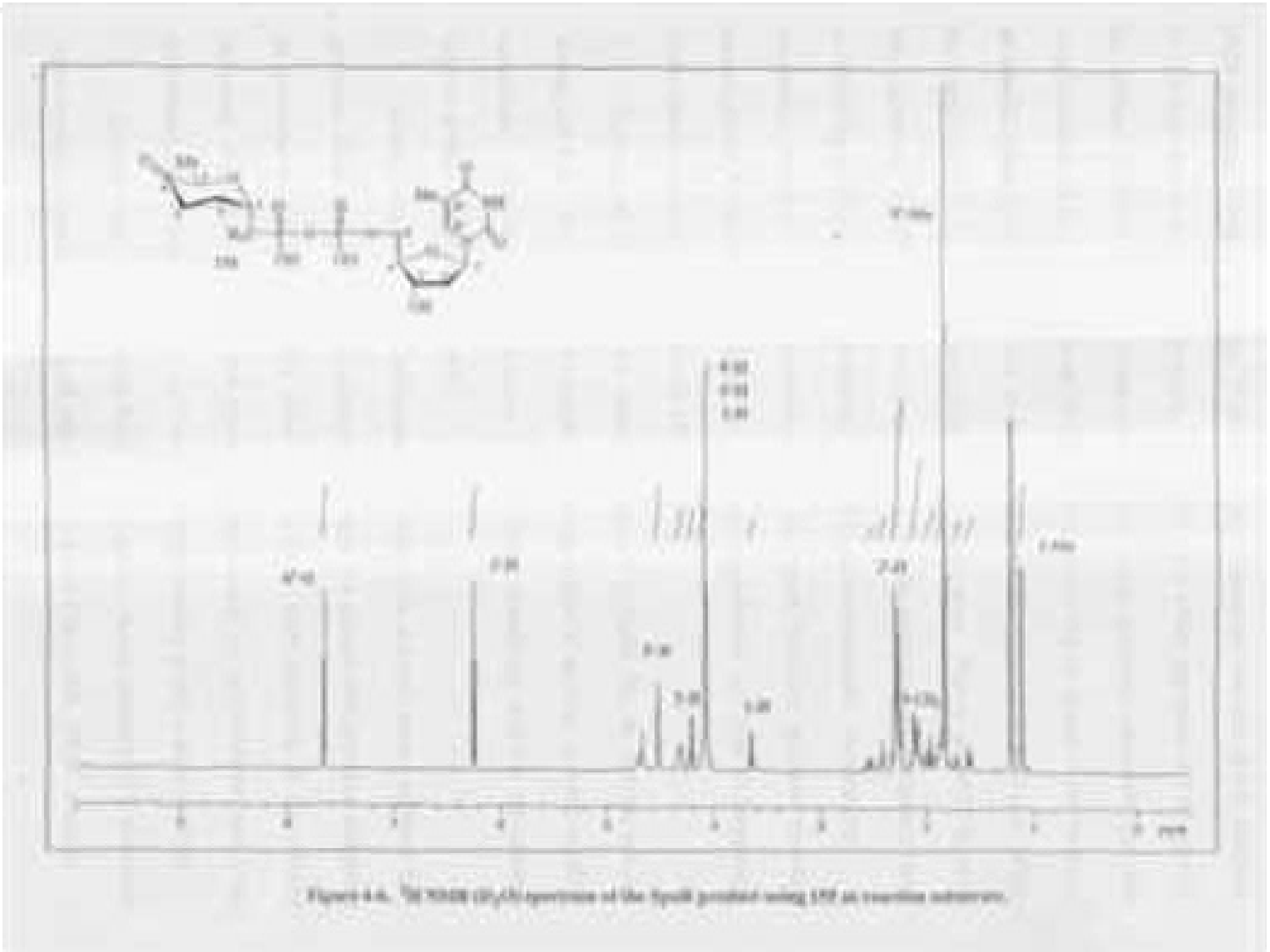
(from amino sugar to keto sugar) in the absence of exogenous PLP at a slow rate and that addition of exogenous PLP accelerated the reaction. When the reconstituted enzymes were used in reactions without exogenous PLP, 34% of the substrate **155** was converted to the product **156** in 10 min while only 5% of **155** was converted in the reaction using the as-isolated enzyme in the same time period. In comparison, 19% of **155** was converted in the SpnR reaction with 50 μM of exogenous PLP in 10 min. The reconstituted enzymes obtained by two reconstitution methods (in the presence or absence of the substrates) did not show a difference in their catalytic efficiency. This correlates well with their similar PLP content.

Isolation and Characterization of the SpnR Product (Figure 4-6). The SpnR reverse reaction product was isolated as a pure compound by gel filtration chromatography when **155** was used as the substrate (Scheme 4-4). The generation of a



Scheme 4-4. Compound 156 was isolated from the large scale SpnR reaction.

keto group at C-4 of the product made it possible that the compound coexisted with its 4-keto hydrated form. This is consistent with the two sets of 5-Me doublet at 1.1 and 1.2



ppm, respectively, and two sets of 5-H quartet at 3.6 and 4.2 ppm, respectively in the ^{1}H NMR spectrum. Moreover, in the ^{31}P NMR spectrum, there are two sets of β-P doublets at 13.9 ppm and 14.8 ppm, respectively. The 4-H signal which appeared at 3.05 ppm in the substrate ^{1}H NMR spectrum disappeared in the product spectrum, consistent with the transformation of the 4-amino group in the substrate to the 4-keto group in the product. In the ^{13}C NMR spectrum, only the hydrated 4-C could be observed at 92.2 ppm while the signal of the C-4 in the carbonyl form could not be seen. This is probably because keto signals in ^{13}C NMR spectra are usually weak, and the amount of the compound used to obtain the spectrum was small, making the keto signal undetectable. In summary, successful isolation and characterization of the SpnR reaction product provided unambiguous evidence that SpnR is an aminotransferase in TDP-D-forosamine biosynthesis, catalyzing conversion of **155** to **156**.

Optimal pH for SpnR. As shown in Figure 4-7, SpnR has an optimal activity around pH 8.5. This enzyme exhibited a sharp reduction of activity in more alkaline conditions. The enzyme can tolerate slightly acidic conditions and has comparable activity as that in neutral conditions.

Amino Acceptors. α-Ketoglutarate was shown to be a better amino acceptor than pyruvate, although pyruvate was a competent alternative (about 50% reaction efficiency of pyruvate compared to α-KG) when the reactions were analyzed in the reverse direction. Since the SpnR catalyzed transamination reaction is a reversible process, it can be inferred that both L-glutamate and L-alanine can be used as amino donors by SpnR.

Activity Assay of SpnR in the Forward Direction. In the presence of L-glutamate and a catalytic amount of PLP, SpnR was able to convert **156** to **155**. HPLC analysis

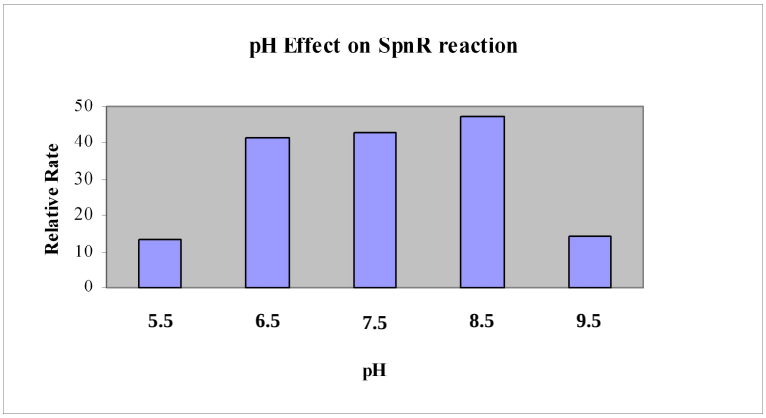


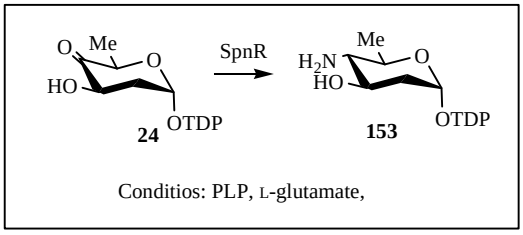
Figure 4-7. pH Effect on the SpnR reaction using 155 as substrate.

showed almost complete conversion from the 4-keto sugar to the 4-amino sugar in 4 h at room temperature. This, again, provided unambiguous evidence that SpnR is indeed the desired aminotransferase that catalyzes the C-4 transamination step in the biosynthesis of TDP-D-forosamine.

SpnR Reaction with the SpnQ Product as the Substrate. After removing all the enzymes in the SpnQ assay mixture, consumption of the SpnQ product accompanied by the concomitant accumulation of a new compound was observed after the addition of SpnR, L-glutamate and PLP in the HPLC analysis. The retention time of the new compound is 9.2 min, which is typical of an amino sugar under the same elution conditions. Since the 4-keto group is the only functional group of the SpnQ product,

TDP-2,3,6-trideoxy-4-keto-glucose (**126**), that is capable of undergoing transamination, this result confirmed that SpnR catalyzes the C-4 transamination of the SpnQ product.

Substrate Specificity of SpnR (Scheme 4-5). In the first reaction, when the SpnN product (**24**) was incubated with SpnR, L-glutamate, and PLP, a new compound formed



Scheme 4-5. SpnR catalyzed reaction with the SpnN product as the substrate.

with a retention time of 14.7 min according to HPLC analysis. In the second reaction, before the SpnQ reaction (condition A including SpnQ, L-glutamate, and PLP) was completed, SpnQ was removed from the reaction mixture, and SpnR and more of L-glutamate and PLP were added. The consumption of the SpnN product continued, and was accompanied by the concomitant continued accumulation of the SpnQ product peak. There were no other peaks formed in the HPLC traces. Thus, the SpnR product formed in the first reaction is same as the SpnQ product formed in the second reaction. The product was identified as TDP-4-amino-2,4,6-trideoxy-glucose (**153**) considering several lines of evidence. First, based on sequence analysis of SpnQ and SpnR, both enzymes contain coenzyme B_6. The common reaction that the two enzymes catalyze in the presence of L-glutamate and PLP should be a transamination reaction. Second, SpnR has been shown to be an efficient aminotransferase in activity assays. The only functional

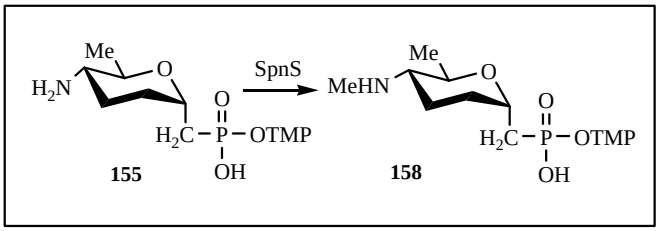
group of the SpnN product (**24**) capable of undergoing an aminotransfer reaction is the 4-keto group. Third, the MS of the isolated compound is 531, which is the same as the calculated mass of TDP-4-amino-2,4,6-trideoxy-glucose (**153**). This result not only provided more supporting evidence that SpnQ is an enzyme that behaves like an aminotransferase under condition A, but also demonstrated the substrate flexibility of SpnR. The enzyme can take both the SpnQ product and the SpnN product as substrates. The presence of the 3-hydroxyl group in the SpnN product does not prevent the compound from being recognized by SpnR. Nonetheless, it should be noted that when the SpnN product was used as the substrate, it took 12 h to achieve 80% substrate conversion even at 37 °C while more than 90% of the SpnQ product could be converted by SpnR in 2 h at room temperature. It is reasonable to speculate that the SpnR catalyzed reaction with the SpnN product (**24**) as substrate is kinetically inefficient because the SpnR active site has evolved to bind the SpnQ 3-deoxygenation product preferentially. If this were not the case, then TDP-D-forosamine could not be obtained since SpnQ requires the 4-keto group to activate the C-3 position for deoxygenation.

Equilibrium Constant. The transamination reaction catalyzed by SpnR is a reversible process. Using **155** and α-KG as substrates, a K_{eq} of 2.6 was obtained from the reaction carried out in 50 mM potassium phosphate buffer (pH 7.5) at 24 °C. The equilibrium slightly favors the direction towards the aminosugar side. This result is consistent with the observation that an excess of α-ketoglutarate was required in order to drive the reverse reaction, from the aminosugar to ketosugar, to completion.

Purification and Characterization of the SpnS Protein (Figure 4-3). The protein was purified to homogeneity after Ni-NTA and Sephacryl S-200 chromatography. Although most of expressed SpnS formed inclusion bodies in the host cells, 50 mg of pure SpnS was nonetheless obtained from a 6 L culture. The absorption spectrum of the purified protein showed no absorbance above 300 nm. Its subunit molecular weight of 28 kDa as revealed by SDS-PAGE correlates well with the predicted value of 28660 Da calculated from the deduced amino acid sequence (including the *C*-terminal His_6-tag). A M*r* of 62.1 kDa, as judged by gel filtration chromatography, indicates that the native SpnS exists as a homodimer. The protein was not stable and precipitated easily, especially after several freeze/thaw cycles. TylM1, the 3-amino-*N*,*N*-dimethyltransferase in the TDP-mycaminose biosynthetic pathway, was reported to be extremely unstable, even requiring the presence of 0.1 mM of exogenous SAM to stabilize the protein during purification. Compared with TylM1, SpnS appears to be a more robust enzyme. No exogenous SAM was included in any buffers used during the purification.

Isolation and Characterization of the SpnS Product (Figure 4-8). The authentic SpnS substrate, TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**), was difficult to obtain. Extensive effort was made to chemically synthesize the compound. However, it turned out that the compound is extremely unstable, the TDP part being easily cleaved. The instability of the compound also made its synthesis through enzyme catalyzed biochemical reactions impractical. Thus, **155** was used in all the SpnS assays as well as the preparative reaction. The preparative reaction product was purified by gel filtration chromatography. Spectroscopic analyses confirmed that it is the monomethylated

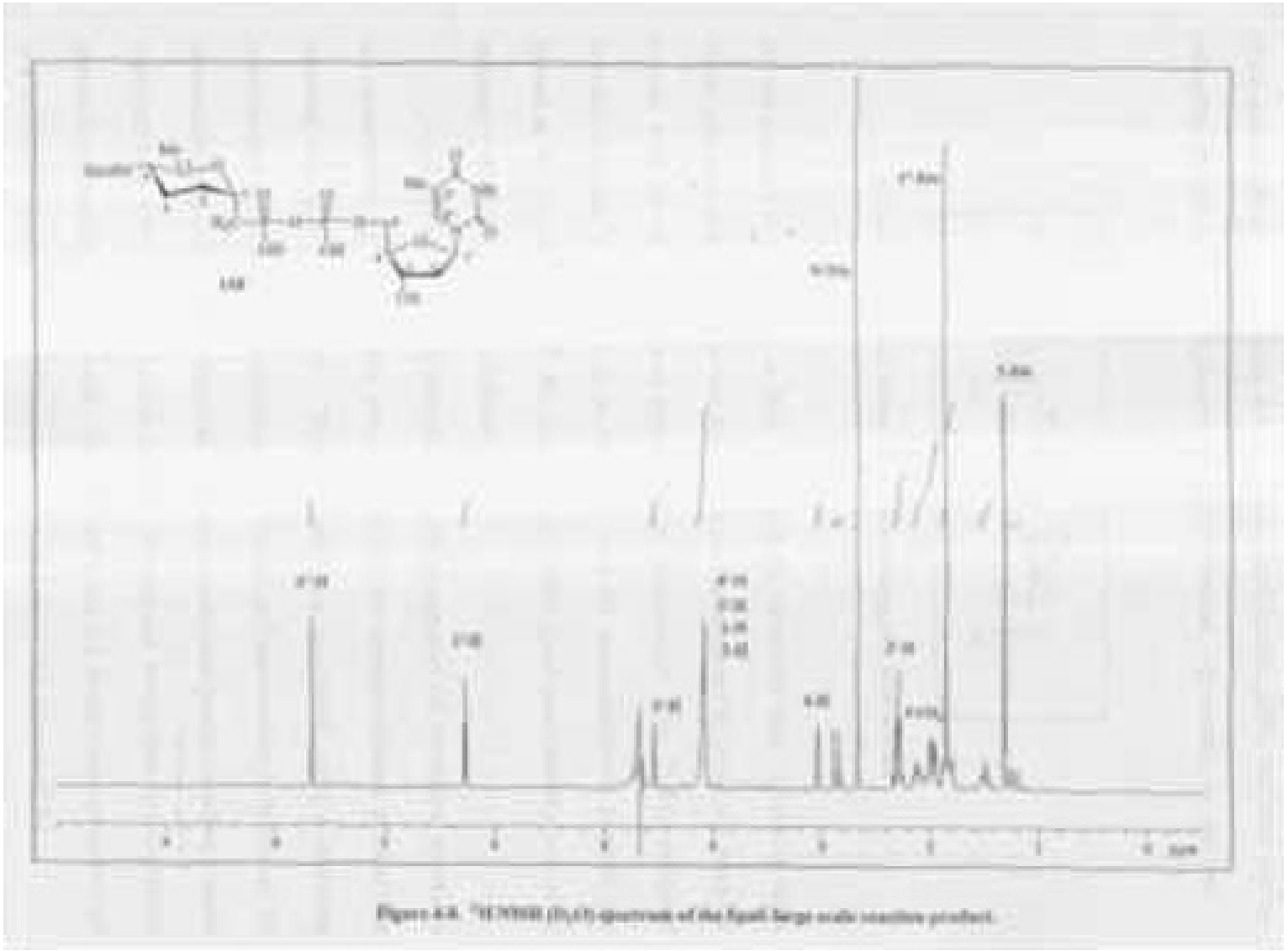
product **158** (Scheme 4-6) . The *N*-methyl group appears at 2.66 ppm as a sharp singlet in the ^{1}H NMR spectrum. Integration of the singlet indicated that it consisted of 3 protons. The *N*-methyl group has a chemical shift of 30.9 ppm in the ^{13}C NMR spectrum. About half of **155** was converted to the monomethylated product (**158**), and there was no noticeable amount of the dimethylated product (**157**) formed. Apparently, the reaction



Scheme 4-6. Compound 158 was isolated from the large scale SpnS reaction.

did not continue to completion to form **157** under the conditions used. However, these results provided direct evidence that the *N*-methylation catalyzed by SpnS is a stepwise process with the monomethylated product as an intermediate. The TylM1 catalyzed dimethylation has been shown to be stepwise, with the monomethylated product also formed as an intermediate.[129,195] But unlike SpnS, TylM1 could quantitatively convert its substrate to the *N*,*N*-dimethylated product.

Reaction Time Course of Methylations. The progress of the reaction as monitored by consumption of the substrate and formation of the monomethylated product was



recorded over 6 h (Figure 4-9). The incubation was conducted using a saturating concentration of SAM. Consumption of the substrate and concomitant formation of the monomethylated product were observed. The monomethylated product reached a steady state one hour after the reaction was initiated and did not accumulate to a significant extent throughout the entire process. On the other hand, the dimethylated product could not be directly observed by HPLC due to peak overlapping. The amount of the

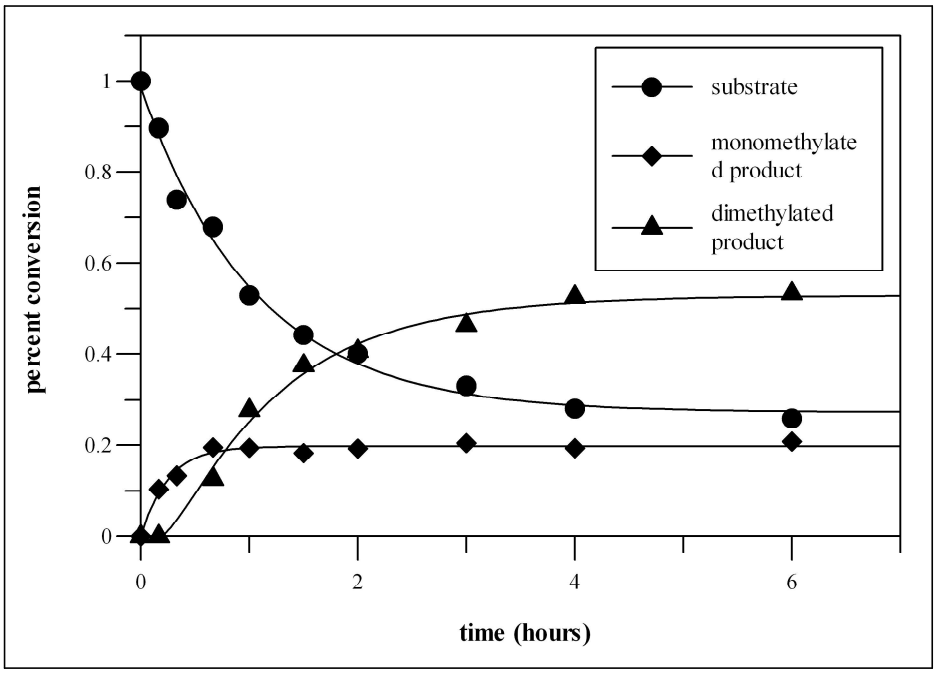


Figure 4-9. Time courses of the SpnS catalyzed N-methyltransfer reactions.

dimethylated product formed was deduced from the difference between the amount of the initial substrate and that of the sum of the remaining substrate and generated monomethylated product. Over time, the amount of the dimethylated product

continuously increased, eventually becoming the major product. Interestingly, only monomethylated product was isolated from the preparative reaction containing the substrate **155**, $MgCl_2$, SAM and SpnS. High concentrations of the substrate and SAM in the preparative reaction might have had an inhibitory effect on SpnS preventing the reaction from proceeding to formation of the dimethylated product. Correspondingly, about half of the amount of the substrate (**155**) that was added initially was recovered even after the reaction had been carried out at 37 °C overnight.

Optimal pH for SpnS. Since neither the monomethylated nor the dimethylated product could be explicitly observed by HPLC using the Dionex column, the pH preference of SpnS was analyzed by the overall consumption of the substrate **155**. SpnS showed an optimal activity around pH 8.5.

Metal Effects. Different di-, and tri-valent cations were examined for their effects on the SpnS reaction rate. When trivalent cations were included in the reaction mixture, no substrate conversion was observed. Among the divalent cations examined, the reaction with Fe^{2+} was most efficient while that with Mg^{2+} was secondary to it. Reactions in the presence of Co^{2+}, Ca^{2+}, and Zn^{2+} did not occur. Considering the fact that Fe^{2+} is easily oxidized by air and Mg^{2+} is the common divalent cation used in NDP containing reactions, Mg^{2+} was included in all subsequent SpnS activity assays and preparative reactions.

Mg^{2+} Concentration Effect. Mg^{2+} is required in the SpnS reaction. In the absence of Mg^{2+}, no substrate consumption was observed. As shown in Figure 4-10, the presence of 2.0 mM of Mg^{2+} increased the reaction rate to its maximum.

Dimethylation of the Monomethylated Product. The isolated monomethylated product, **158**, was incubated with SpnS along with other necessary components. In this reaction, the competition for SAM between the monomethylation and dimethylation reactions was removed. From the MS analysis of the product mixture, the presence of m/z of 540 in ESI-MS $(M-H)^-$ and m/z of 542 in ESI-MS $(M+H)^+$ was observed. The dimethylated product has a molecular weight of 541. The MS results indicate that SpnS is able to catalyze the methyltransfer reaction using the monomethylated product as the substrate, confirming the stepwise mechanism of the SpnS reaction.

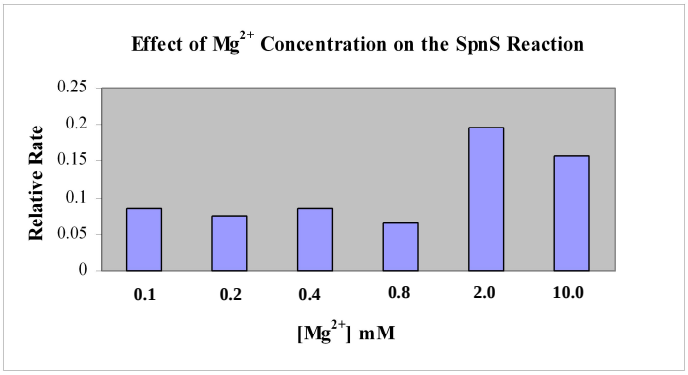


Figure 4-10. Effect of Mg^{2+} concentration on the relative rate of the SpnS reaction.

4. CONCLUSIONS

The results from this chapter provide unambiguous evidence that SpnR and SpnS are the aminotransferase and methyltransferase in the biosynthetic pathway of TDP-D-forosamine, catalyzing conversion of TDP-2,3,6-trideoxy-4-keto-glucose (**126)** to TDP-4-amino-2,3,4,6-tetradeoxy-glucose (**127**) and TDP-4-amino-2,3,4,6-tetradeoxy-glucose

(**127)** to TDP-D-forosamine (**128**), respectively. The detailed study carried out on the PLP-containing SpnR will enrich our understanding of the C-4 aminotransferases that occur in several other deoxysugar biosynthetic pathways, and will also pave the way to a better understanding of the evolutionary differences between C-3 aminotransferases and C-4 aminotransferases. The SpnS catalyzed reaction is another example of a SAM-dependent methylation reaction. The aerobic reaction conditions used eliminate the possibility of SAM radical involvement. So far, all the methylation reactions that occur on the sugar components of antibiotics that have been characterized use SAM as the methyl donor. More importantly, the investigation in this chapter concludes the biosynthetic studies of TDP-D-forosamine. The biosynthetic pathway of TDP-D-forosamine has been unequivocally established as a series of reactions catalyzed by SpnO, SpnN, SpnQ, SpnR and SpnS in sequence.

REFERENCES

1) Lindberg, B. "Components of bacterial polysaccharides" *Adv. Carbohydr. Chem. Biochem*. **1990**, 48, 279.

2) Weymouth-Wilson, A. C. "The role of carbohydrates in biologically active natural products" *Nat. Prod. Rep*. **1997**, 14, 99.

3) Kisailus, E. C. *Glycoconjugates: Composition, Structure, and Function*; Marcel Dekker Inc.: New York, 1992.

4) Montreuil, J.; Vleigenthart, J. F. G.; Schachter, H. *Glycoproteins*; Elsevier: Amsterdam, 1995.

5) Wiegandt, H. *Glycolipids*; Elsevier: Amsterdam, 1985.

6) Varki, A. *Glycobiology*. **1997**, 3, 97.

7) Montreuil, J.; Vleigenthart, J. F. G.; Schachter, H. *Glycoproteins and Disease*; Elsevier: Amsterdam, 1996.

8) Meldgaard, P.; Holmes, E. H.; Bennett, E.P.; Clausen, H.; Zeuthen, J.; Wolf, H.; Orntoft, T.F. "Blood group ABO-related glycosylation of urothelial cell lines: immunocytological, enzymatic, and genetic characterization" *Cancer Res*. **1994,** 54 (9), 2440.

9) Streiff, M. B.; Segal, J.; Grossman, S. A.; Kickler, T. S.; Weir, E. G. "ABO blood group is a potent risk factor for venous thromboembolism in patients with malignant gliomas" *Cancer* **2004**, 100 (8), 1717.

10) Wu, A. M. *The Molecular Immunology of Complex Carbohydrates-2*; Kluwer Academic: New York, 2001.

11) Wu, A. M. “Carbohydrate structural units in glycoproteins and polysaccharides as important ligands for Gal and GalNAc reactive lectins” *J. Biomed. Sci*. **2003**, 10(6 Pt 2), 676.

12) Wong, C. H. *Carbohydrate-Based Drug Discovery*; Weinheim: Cambridge, 2003.

13) Hanessian, s.; Haskell, T. H. *The Carbohydrates: Chemistry and Biochemistry*; Academic Press: New York, 1970, vol IIA, 139.

14) Courtois, J. E.; Percheron, F. *The Carbohydrates: Chemistry and Biochemistry*; Academic Press: New York, 1970, vol IIA, 213.

15) Williams, N. R.; Wander, J. D. *The Carbohydrates: Chemistry and Biochemistry*; Academic Press: New York, 1980, vol IB, 761.

16) Schaffer, R. *The Carbohydrates: Chemistry and Biochemistry*; Academic Press: New York, 1972, vol 1A, 69.

17) Molhoj, M.; Verma, R.; Reiter, W. D. “The biosynthesis of the branched-chain sugar d-apiose in plants: functional cloning and characterization of a UDP-d-apiose/UDP-d-xylose synthase from Arabidopsis” *Plant J.* **2003**, 35(6), 693.

18) Liu, H. M.; Zhang, F.; Zhang, J. “Stereoselective synthesis of methyl 4,6-O-benzylidene-2-C-methoxycarbonylmethyl-alpha-D-ribo-hexopyranosid-3-ulose, and its X-ray crystallographic analysis” *Carbohydr. Res***. 2001**, 334(4), 323.

19) Zhao, S.; Petrus, L.; Serianni, A. S.; “1-Deoxy-D-xylulose: synthesis based on molybdate-catalyzed rearrangement of a branched-chain aldotetrose” *Org. Lett.* **2001**, 3(24), 3819.

20) Hawkins, L. D; Christ, W. J.; Rossignol, D. P. “Inhibition of endotoxin response by synthetic TLR4 antagonists” *Curr. Top. Med. Chem*. **2004**, 4(11), 1147.

21) Trent, M. S. “Biosynthesis, transport, and modification of lipid A” *Biochem. Cell. Biol*. **2004**, 82(1), 71.

22) Cavaillon, J. M.; Adrie, C.; Fitting, C.; Adib-Conquy, M. “Endotoxin tolerance: is there a clinical relevance?” *J. Endotoxin Res*. **2003**, 9(2), 101.

23) Yethon, J. A.; Whitfield, C. “Lipopolysaccharide as a target for the development of novel therapeutics in gram-negative bacteria” *Curr. Drug Targets Infect. Disord*. **2001**, 1(2), 91.

24) Trefzer, A.; Salas, J. A.; Bechthold, A. “Genes and enzymes involved in deoxysugar biosynthesis in bacteria” *Nat. Prod. Rep***1999** , 16, 283.

25) Rietschel, E. T.; Kirikae, T.; Schade, F. U.; Mamat, U.; Schmidt, G.; Loppnow, H.; Ulmer, A. J.; Zahringer, U.; Seydel, U.; Di Padova, F. “Bacterial endotoxin: molecular relationships of structure to activity and function” *FASEB J*. **1994**, 8(2), 217.

26) Bosman, F. T.; Stamenkovic, I. “Functional structure and composition of the extracellular matrix” *J. Pathol*. **2003**, 200(4), 423.

27) Anthony, V.; Skach, W. R. “Molecular mechanism of P-glycoprotein assembly into cellular membranes” *Curr. Protein Pept. Sci* **2002** , 3(5), 485.

28) Skurnik, M.; Bengoechea, J. A. “The biosynthesis and biological role of lipopolysaccharide O-antigens of pathogenic Yersiniae” *Carbohydr. Res*. **2003**, 338(23), 2521.

29) Lerouge, I.; Vanderleyden, J. “O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions” *FEMS Microbiol. Rev*. **2002**, 26(1), 17.

30) Alavi, A.; Axford, J. S. *Glycoimmunology*; Plenum Press: New York, 1995.

31) Caroff, M.; Karibian, D. “Structure of bacterial lipopolysaccharides” *Carbohydr. Res*. **2003**, 338(23), 2431.

32) Bengoechea JA, Najdenski H, Skurnik M. “Lipopolysaccharide O antigen status of Yersinia enterocolitica O: 8 is essential for virulence and absence of O antigen affects the expression of other Yersinia virulence factors” *Mol. Microbiol*. **2004**, 52(2), 451.

33) Holme, T.; Rahman, M.; Jansson, P. E.; Widmalm. G. “The lipopolysaccharide of moraxella catarrhalis structural relationships and antigenic properties” *Eur. J. Biochem*. **1999**, 265(2), 524.

34) Penner, J. L.; Aspinall, G. O. “Diversity of lipopolysaccharide structures in Campylobacter jejuni” *J. Infect. Dis.* **1997**, 176, Suppl 2, S135-8.

35) Moran, A. P.; Prendergast, M. M.; Appelmelk, B. J. “Molecular mimicry of host structures by bacterial lipopolysaccharides and its contribution to disease” *FEMS Immunol. Med. Microbiol*. **1996**, 16(2), 105.

36) De Castro, C.; De Castro, O.; Molinaro, A.; Parrilli, M. “Structural determination of the O-chain polysaccharide from Agrobacterium tumefaciens, strain DSM 30205” *Eur. J. Biochem*. **2002**, 269(12), 2885.

37) Kannenberg, E. L.; Carlson, R. W. “Lipid A and O-chain modifications cause Rhizobium lipopolysaccharides to become hydrophobic during bacteroid development” *Mol. Microbiol*. **2001,** 39(2), 379.

38) Kennedy, J. F.; White, C. A. *Bioactive Carbohydrates in Chemistry, Biochemistry, and Biology*; Ellis Horwood Publishers: Chichester, 1983.

39) Liu, H. W.; Thorson, J. S. “Pathways and mechanisms in the biogenesis of novel deoxysugars by bacteria” *Annu. Rev. Microbiol*. **1994**, 48, 223.

40) Nissen, P.; Hansen, J.; Ban N.; Moore, P. B.; Steitz, T. A. “The structural basis of ribosome activity in peptide bond synthesis’ *Science* **2000**, 289, 920.

41) Schlunzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. Structural basis for the interaction of antibiotics with the peptidyl transferase center in eubacteria” *Nature* **2001**, 413, 814.

42) Zhou, W.; Gumina, G.; Chong, Y.; Wang, J.; Schinazi, R. F.; Chu, C. K. “Synthesis, structure-activity relationships, and drug resistance of beta-d-3'-fluoro-2',3'-unsaturated nucleosides as anti-HIV Agents” *J. Med. Chem*. **2004**, 47(13), 3399.

43) Kim, C.; Haddad, J.; Vakulenko, S. B.; Meroueh, S. O.; Wu, Y.; Yan, H.; Mobashery, S. “Fluorinated aminoglycosides and their mechanistic implication for aminoglycoside 3'-phosphotransferases from Gram-negative bacteria” *Biochemistry* **2004**, 43(9), 2373.

44) Kobayashi, Y.; Ohgami, T.; Ohtsuki, K.; Tsuchiya, T. “Synthesis of kanamycin A analogs containing a 6-amino-6-deoxyglycofuranose moiety” *Carbohydr. Res*. **2000**, 329(2), 325.

45) Mendez, C.; Salas, J. A. “Altering the glycosylaion pattern of bioactive compounds” *Trends Biotechnol*. **2001**, 19, 449.

46) Thorson, J. S.; Hoster, T. J.; Jiang, J.; Biggins, J. B.; Ahlert, J. “Nature’s carbohydrate chemists: the enzymatic glycosylation of bioactive bacterial metabolites” *Curr. Org. Chem*. **2001**, 5, 139.

47) Salah-Bey, K. “Targeted gene inactivation for the elucidation of deoxysugar biosynthesis in the erythromycin producer *Saccharopolyspora erythraea*” *Mol. Gen. Genet*. **1998**, 257, 542.

48) Gonzalez, A. “The *mtm VUC* genes of the mithramycin gene cluster in *Streptomyces argillaceus* are involved in the biosynthesis of the sugar moieties” *Mol. Gen. Genet*. **2001**, 264, 827.

49) Zhao, L.; Que, N. L. S.; Xue, Y. Q.; Sherman, D. H.; Liu, H. W. “Mechanistic studies of desosamine biosynthesis: C-4 deoxygenation precedes C-3 transamination” *J. Am. Chem. Soc*. **1998**, 120, 12159.

50) Zhao, L.; Ahlert, J.; Xue, Y. Q.; Thorson, J. S.; Sherman, D. H.; Liu, H. W. “Engineering a methymycin/pikromycin-calicheamicin hybrid: construction of two new macrolides carrying a designed sugar moiety” *J. Am. Chem. Soc*. **1999**, 121, 9881.

51) Borisova, S.A; Zhao, L.; Sherman, D. H.; Liu, H. W. “Biosynthesis of desosamine: construction of a new macrolide carrying a genetically designed sugar moiety” *Org. Chem*. **1999**, 15, 133.

52) He, X. M.; Liu, H. W. “Formation of unusual sugars: mechanistic studies and biosynthetic applications” *Annu. Rev. Biochem*. **2002**, 71, 701.

53) He, X. M.; Liu, H. W. “Mechanisms of enzymatic C-O bond cleavages in deoxyhexose biosynthesis” *Curr. Opin. Chem. Biol*. **2002**, 6, 590.

54) Johnson, D. A.; Liu, H. W. “Mechanisms and pathways from recent deoxysugar biosynthesis research” *Curr. Opin. Chem. Biol.* **1998**, 2(5), 642.

55) Lei. Y.; Ploux, O.; Liu, H. W. “Mechanistic studies on CDP-6-deoxy-L-threo-D-glycero-4-hexulose 3-dehydrase identification of His-220 as the active-site base by chemical modification and site-directed mutagenesis” *Biochemistry* **1995,** 34(14), 4643.

56) Ploux, O.; Leik, Y.; Vatanen, K.; Liu, H. W. “Mechanistic studies on CDP-6-deoxy-$\Delta^{3,4}$-glucoseen reductase: the role of cysteine residues in catalysis as probed by chemical modification and site-directed mutagenesis” *Biochemistry* **1995**, 34(13), 4159.

57) Lo, S. F.; Miller, V. P.; Lei, Y.; Thorson, J. S.; Liu, H. W.; Schottel, J. L. “CDP-6-deoxy-delta 3,4-glucoseen reductase from Yersinia pseudotuberculosis: enzyme purification and characterization of the cloned gene” *J. Bacteriol*. **1994**, 176(2), 460.

58) Gassner, G. T.; Johnson, D. A.; Liu, H. W.; Ballou, D. P. “Kinetics of the reductive half-reaction of the iron-sulfur flavoenzyme CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase reductase” *Biochemistry* **1996**, 35, 7752.

59) Johnson, D. A.; Gassner, G. T.; Bandarian, V.; Ruzicka, F. J.; Ballou, D. P.; Reed, G. H.; Liu, H. W. “Kinetic characterization of an organic radical in the ascarylose biosynthetic pathway” *Biochemistry* **1996**, 35, 15846.

60) Burns, K. D.; Pieper, P. A.; Liu, H. W.; Stankovich, M. T. “CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase (E1) and CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase reductase (E3): two important enzymes involved in the biosynthesis of ascarylose” *Biochemistry* **1996**, 35, 7879.

61) Chen, X. M. H.; Polux, O.; Liu, H. W. “Biosynthesis of 3, 6-Dideoxyhexoses: In Vivo and in vitro evidence for protein-protein interaction between CDP-6-deoxy-L-*threo*-D-*glycero*-4-hexulose-3-dehydrase (E1) and its reductase (E3)” *Biochemistry* **1996**, 35, 16412.

62) Miller, V. P.; Thorson, J. S.; Ploux, O.; Lo, S. F.; Liu, H. W. “Cofactor characterization and mechanistic studies of CDP-6-deoxy-$\Delta^{3,4}$-glucoseen reductase : exploration into a novel enzymatic C-O bond cleavage event” *Biochemistry* **1993**, 32, 11934.

63) Weigel, T. M.; Miller, V. P.; Liu, H. W. “Mechanistic and stereochemical studies of a unique dehydration catalyzed by CDP-4-keto-6-deoxy-D-glucose-3-dehydrase: a pyridoxamine 5’-phosphate dependent enzyme isolated from *Yersinia pseudotuberculosis*” *Biochemistry* **1992**, 31, 2140.

64) Weigel, T. M.; Liu, L.; Liu, H. W. “Mechanistic studies of the biosynthesis of 3, 6-dideoxyhexoses in *Yersinia pseudotuberculosis*: purification and characterization of CDP-4-keto-6-deoxy-D-glucose-3-dehydrase” *Biochemistry* **1992**, 31, 2129.

65) Thorson, J. S.; Liu, H. W. “Characterization of the first PMP-dependent iron-sulfur-containing enzyme which is essential for the biosynthesis of 3, 6-dideoxyhexoses” *J. Am. Chem. Soc*. **1993**, 115, 7539.

66) Thorson, J. S.; Liu, H. W. “Coenzyme B_6 as a redox cofactor: a new role for an old coenzyme?”*J. Am. Chem. Soc* . **1993**, 115, 12177.

67) Chang, C. W. T.; Chen, X. H.; Liu, H. W. “CDP-6-deoxy-6, 6-difluoro-D-glucose: a mechanism-based inhibitor for CDP-D-glucose 4, 6-dehydrase” *J. Am. Chem. Soc*. **1998**, 120, 9698.

68) Chang, C. W. T.; Johnson, D. A.; Bandarian, V.; Zhou, H.; LoBrutto, R.; Reed, G. H.; Liu, H. W. “Characterization of a unique coenzyme B_6 radical in the ascarylose biosynthetic pathway” *J. Am. Chem. Soc*. **2000**, 122, 4239.

69) Beyer, N.; Alam, J.; Hallis, T. M.; Guo, Z.; Liu, H. W. “The biosynthesis of GDP-L-colitose: C-3 deoxygenation is catalyzed by a unique coenzyme B_6-dependent enzyme” *J. Am. Chem. Soc*. **2003**, 125, 5584.

70) Draeger, G.; Park, S. H.; Floss, H. G. “Mechanism of the 2-deoxygenation step in the biosynthesis of the deoxyhexose moieties of the antibiotics granaticin and oleandomycin” *J. Am. Chem. Soc*. **1999**, 121, 2611.

71) Chen, H.; Agnihotri, G.; Guo, Z.; Que, N. L. S.; Chen, X. H.; Liu, H. W. "Biosynthesis of mycarose: isolation and characterization of enzymes involved in the C-2 deoxygenation" *J. Am. Chem. Soc*. **1999**, 121, 8124.

72) Zhao, L.; Borisova, S.; Yeung, S. M.; Liu, H. W. "Study of C-4 deoxygenation in the biosynthesis of desosamine: evidence implicating a novel mechanism" *J. Am. Chem. Soc*. **2001**, 123, 7909.

73) Zhang, Q.; Liu, H. W. "Chemical synthesis of UDP-β-L-arabinofuranose and its turnover to UDP-β-L-arabinopyranose by UDP-galactopyranose mutase" *Bioorg. Med. Chem. Lett.* **2001**, 11, 145.

74) Huang, Z.; Zhang, Q.; Liu, H. W. "Reconstitution of UDP-galactopyranose mutase with 1-deaza-FAD and 5-deaza-FAD: analysis and mechanistic implications" *Bioorg. Chem.* **2003**, 31, 494.

75) Nedal, A.; Zotchev, S. B. "Biosynthesis of deoxyaminosugars in antibiotic-producing bacteria" *Appl. Microbiol. Biotechnol*. **2004**, 64, 7.

76) Kolberg, M.; Strand, K. R.; Graff, P.; Andersson, K, K. "Structure, function, and mechanism of ribonucleotide reductases" *Biochim. Biophys. Acta*. **2004**, 1699(1-2), 1.

77) Follmann, H. "Deoxyribonucleotides: the unusual chemistry and biochemistry of DNA precursors" *Chem. Soc. Rev*. **2004**, 33(4), 225.

78) Cox, M. M. "Better chemistry for better survival, through regulation" *Cell* **2003**, 112(3), 286.

79) Hans, M.; Homung, A.; Dziarnowski, A.; Cane, D. E.; Khosla, C. “Mechanistic analysis of acyl transferase domain exchange in polyketide synthase modules” *J. Am. Chem. Soc*. **2003**, 125, 5366.

80) Chen, H.; Guo, Z.; Liu, H. W. “Expression, purification, and characterization of TylM1, an *N*, *N*-dimethyltransferase involved in the biosynthesis of mycaminose” *J. Am. Chem. Soc*. **1998**, 120, 9951.

81) Chen, H.; Yeung, S. M.; Que, N. L. S.; Muller, T.; Schmidt, R. R.; Liu, H. W. “Expression, purification, and characterization of TylB, an aminotransferase involved in the biosynthesis of mycaminose” *J. Am. Chem. Soc*. **1999**, 121, 7166.

82) Walsh, C.; Meyers, C. L. F.; Losey, H. C. “Antibiotic glycosyltransferases: antibiotic maturation and prospects for reprogramming” *J. Med. Chem.* **2003**, 46(16), 3425.

83) Schonfeld, W.; Kirst, H. A. *Macrolide antibiotics*; Birkhauser Verlag: Basel, 2002.

84) Metsa-Ketela, M.; Palmu, K.; Kunnari, T.; Ylihonko, K.; Mantsala, P. “Engineering anthracycline biosynthesis toward angucyclines” *Antimicrob. Agents Chemother*. **2003**, 47(4), 1291.

85) Raty, K.; Kantola, J.; Hautala, A.; Hakala, J.; Ylihonko, K.; Mantsala, P. “Cloning and characterization of Streptomyces galilaeus aclacinomycins polyketide synthase (PKS) cluster” *Gene* **2002**, 293(1-2), 115.

86) Rawlings, B. J. “Type I polyketide biosynthesis in bacteria (Part A--erythromycin biosynthesis)” *Nat. Prod. Rep*. **2001**, 18(2), 190

87) Walsh, C. T.; O'Connor, S. E.; Schneider, T. L. “Polyketide-nonribosomal peptide epothilone antitumor agents: the EpoA, B, C subunits” *Ind. Microbiol. Biotechnol*. **2003**, 30(8), 448.

88) Walsh, C. T. “Polyketide and nonribosomal peptide antibiotics: modularity and versatility” *Science* **2004**, 303, 1805.

89) Xu, H.; Heide, L.; Li, S. M. “New aminocoumarin antibiotics formed by a combined mutational and chemoenzymatic approach utilizing the carbamoyltransferase NovN” *Chem. Biol*. **2004**, 11(5), 655.

90) Eustaquio, A. S.; Luft, T.; Wang, Z. X.; Gust, B.; Chater, K. F.; Li, S. M.; Heide, L. “Novobiocin biosynthesis: inactivation of the putative regulatory gene novE and heterologous expression of genes involved in aminocoumarin ring formation” *Arch. Microbiol*. **2003**, 180(1), 25.

91) Walsh, C. T.; Gehring, A. M.; Weinreb, P. H.; Quadri, L. E.; Flugel, R. S. “Post-translational modification of polyketide and nonribosomal peptide synthases” *Curr. Opin. Chem. Biol*. **1997**, 1(3), 309.

92) Walsh, C. T. *Antibiotics: Actions, Origins, Resistance*; ASM press: Washington, D.C. 2003.

93) Keating, T. A.; Walsh, C. T. “Initiation, elongation, and termination strategies in polyketide and polypeptide antibiotic biosynthesis” *Curr. Opin. Chem. Biol*. **1999**, 3(5), 598.

94) Madduri, K.; Waldron, C.; Merlo, D. J. “Rhamnose biosynthesis pathway supplies precursors for primary and secondary metabolism in *Saccharopolyspora spinosa*” *J. Bacteriol*. **2001**, 183(19), 5632.

95) Giraud, M. F.; Naismith, J. H. “The rhamnose pathway” *Curr. Opin. Struct. Biol*. **2000**, 10(6), 687.

96) Barber, G. A. “The synthesis of guanosine 5'-diphosphate D-rhamnose by enzymes of a higher plant” *Biochim. Biophys. Acta*. **1968**, 165(1), 68.

97) Eustaquio, A. S.; Luft, T.; Wang, Z. X.; Gust, B.; Chater, K. F.; Li, S. M.; Heide, L. “Novobiocin biosynthesis: inactivation of the putative regulatory gene *nov*E and heterologous expression of genes involved in aminocoumarin ring formation” *Arch. Microbiol*. **2003**, 180, 25.

98) Galm, U.; Schimana, J.; Fiedler, H. P.; Schmidt, J.; Li, S. M.; Heide, L. “Cloning and analysis of the simocyclinone biosynthetic gene cluster of *Streptomyces antibioticus Tu* 6040” *Arch. Microbiol*. **2002**, 178(2), 102.

99) Torkkell, S.; Kunnari, T.; Palmu, K.; Mantsala, P.; Hakala, J.; Ylihonko, K. “The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures” *Mol. Genet. Genomics*. **2001**, 266, 276.

100) Menendez, N.; Nur-e-Alam, M.; Brana, A. F.; Rohr, J.; Salas, J. A.; Mendez, C. “Biosynthesis of the antitumor chromomycin A3 in *Streptomyces griseus*: analysis of the gene cluster and rational design of novel chromomycin analogs” *Chem. Biol*. **2004**, 11(1), 21.

101) Trefzer, A.; Pelzer, S.; Schimana, J.; Stockert, S.; Bihlmaier, C.; Fiedler, H. P.; Welzel, K.; Vente, A.; Bechthold, A. “Biosynthetic gene cluster of simocyclinone, a natural multihybrid antibiotic” *Antimicrob. Agents Chemother*. **2002**, 46, 1174.

102) Hoffmeister, D.; Lchinose, K.; Domann, S.; Faust, B.; Trefzer, A.; Drager, G.; Kirschning, A.; Fischer, C.; Kunzel, E.; Bearden, D. W.; Rohr, J.; Bechthold, A. “The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster” *Chem. Biol*. **2000**, 7, 821.

103) Gaisser, S.; Bohm, G. A.; Cortes, J.; Leadlay, P. F. “Analysis of seven genes from the eryAI- eryK region of the erythromycin biosynthetic gene cluster in *Saccharopolyspora erythraea*” *Mol. Gen. Genet*. **1997**, 256, 239.

104) Gonzalez, A.; Remsing, L. L.; Lombo, F.; Fernandez, M. J.; Prado, L.; Brana, A. F.; Kunzel, E.; Rohr, J.; Mendez, C.; Salas, J. A. “The mtmVUC genes of the mithramycin gene cluster in Streptomyces argillaceus are involved in the biosynthesis of the sugar moieties” *Mol. Gen. Genet*. **2001**, 264(6), 827.

105) Yoon, Y. J.; Kim, E. S.; Hwang, Y. S.; Choi, C. Y. “Avermectin: biochemical and molecular basis of its biosynthesis and regulation” *Appl. Microbiol. Biotechnol*. **2004**, 63, 626.

106) Aguirrezabalaga, I.; Olano, C.; Allende, N.; Rodriguez, L.; Brana, A. F.; Mendez, C.; Salas, J. “Identification and expression of genes involved in biosynthesis of L-

oleandrose and its intermediate L-olivose in the oleandomycin producer Streptomyces antibioticus" *Antimicrob. Agents Chemother*. **2000**, 44, 1266.

107) Rodriguez, L.; Rodriguez, D.; Olano, C.; Brana, A. F.; Mendez, C.; Salas, J. "Functional analysis of OleY L-oleandrosyl 3-*O*-methyltransferase of the oleandomycin biosynthetic pathway in *Streptomyces antibioticus*" *J. Bacteriol.* **2001**, 183, 5358.

108) Kessler, A. C.; Haase, A.; Reeves, P. R. "Molecular analysis of the 3, 6-dideoxyhexose pathway genes of Yersinia pseudotuberculosis serogroup IIA" *J. Bacteriol.* **1993**, 175, 1412.

109) Cundliffe, E.; Bate, N.; Butler, A.; Fish, S.; Gandecha, A.; Merson-Davies, L. "The tylosin-biosynthetic genes of Streptomyces fradiae" *Antonie. Van. Leeuwenhoek*. **2001**, 79(3-4), 229.

110) Aparicio, J. F.; Caffrey, P.; Gil, J. A.; Zotchev, S. B. "Polyene antibiotic biosynthesis gene clusters" *Appl. Microbiol. Biotechnol*. **2003**, 61, 179.

111) Zotchev, S. B. "Polyene macrolide antibiotics and their applications in human therapy" *Curr. Med. Chem.* **2003**, 10, 211.

112) Hoffmeister, D.; Ichinose, K.; Domann, S.; Faust, B.; Trefzer, A.; Drager, G.; Kirschning, A.; Fischer, C.; Kunzel, E.; Bearden, D. W.; Rohr, J.; Bechthold, A. "The NDP-sugar co-substrate concentration and the enzyme expression level influence the substrate specificity of glycosyltransferases: cloning and characterization of deoxysugar biosynthetic genes of the urdamycin biosynthetic gene cluster" *Chem. Biol*. **2000**, 7, 821.

113) Westrich, L.; Domann, S.; Faust, B.; Bedford, D.; Hopwood, D. A.; Bechthold, A. "Cloning and characterization of a gene cluster from *Streptomyces cyanogenus* S136 probably involved in landomycin biosynthesis" *FEMS Microbiol. Lett.* **1999**, 170, 381.

114) Sun, Y.; Zhou, X.; Dong, H.; Tu, G.; Wang, M.; Wang, B.; Deng, Z. "A complete gene cluster from *Streptomyces nanchangensis* NS 3226 encoding biosynthesis of the polyether ionophore nanchangmycin" *Chem. Biol.* **2003**, 10, 431.

115) Raty, K.; Kunnari, T.; Hakala, J.; Mantsala, P.; Ylihonko, K. "A gene cluster from *Streptomyces galilaeus* involved in glycosylation of aclarubicin" *Mol. Gen. Genet.* **2000**, 264, 164.

116) Torkkell, S.; Kunnari, T.; Palmu, K.; Mantsala, P.; Hakala, J.; Ylihonko, K. "The entire nogalamycin biosynthetic gene cluster of *Streptomyces nogalater*: characterization of a 20-kb DNA region and generation of hybrid structures" *Mol. Genet. Genomics*. **2003**, 266, 276.

117) Van Wageningen, A. M.; Kirkpatrick, P. N.; Williams, D. H.; Harris, B. R.; Kershaw, J. K.; Lennard, N. J.; Jones, M.; Jones, S. J. M.; Solenberg, P. J. "Sequencing and analysis of genes involved in the biosynthesis of a vancomycin group antibiotic" *Chem. Biol.* **1998**, 5(3), 155

118) Hallis, M.; Liu, H. W. "Learning nature's strategies for making deoxy sugars: pathways, mechanisms, and combinatorial applications" *Acc. Chem. Res*. **1999**, 32, 579.

119) He, X.; Agnihotri, G. Liu, H. W. “Novel enzymatic mechanisms in carbohydrate metabolism” *Chem. Rev*. **2000**, 100, 4615.

120) Snipes, C. E.; Brillinger, G. U.; Sellers, L.; Mascaro, L.; Floss, H. G. “Stereochemistry of the dTDP-glucose oxidoreductase reaction” *J. Biol. Chem*. **1977**, 252, 8113.

121) Russell, R. N.; Liu, H. W. “Stereochemical and mechanistic studies of CDP-D-glucose oxidoreductase isolated from *Yersinia pseudotuberculosis*” *J. Am. Chem. Soc*. **1991**, 113, 7777.

122) Yu, Y.; Russell, R. N.; Thorson, J. S.; Liu, L. D.; Liu, H. W. “Mechanistic studies of the biosynthesis of 3, 6-dideoxyhexoses in *Yersinia pseudotuberculosis*: purification and stereochemical analysis of CDP-D-glucose oxidoreductase” *J. Biol. Chem*. **1992**, 267, 5868.

123) Hegeman, A.D.; Gross, J. W.; Frey, P. A. “Concerted and stepwise dehydration mechanisms observed in wild-type and mutated *Escherichia coli* dTDP-glucose 4,6-dehydratase” *Biochemistry* **2002**, 41, 2797.

124) Bauer, A. J.; Rayment, I.; Frey, P. A.; Holden, H. M. “The molecular structure of UDP galactose 4-epimerase from *Escherichia coli* determined at 2.5 Å resolution” *Prot. Struct. Funct. Genet*. **1992**, 12, 372.

125) Hegeman, A. D.; Gross, J. W.; Frey, P. A. “Probing catalysis by *Escherichia coli* dTDP-glucose-4, 6-dehydratase: identification and preliminary characterization of functional amino acid residues at the active site” *Biochemistry* **2001**, 40, 6598.

126) Gerratana, B.; Cleland, W. W.; Frey, P. A. “Mechanistic roles of Thr134, Tyr160, and Lys164 in the reaction catalyzed by dTDP-glucose 4, 6-dehydratase” *Biochemistry* **2001**, 40, 9187.

127) Ichinose, K.; Ozawa, M.; Itou, K.; Kunieda, K.; Ebizuka, Y. “Cloning, sequencing and heterologous expression of the medermycin biosynthetic gene cluster of Streptomyces sp. AM-7161: towards comparative analysis of the benzoisochromanequinone gene clusters” *Microbiology* **2003**, 149, 1633.

128) Gross, J. W.; Hegeman, A. D.; Gerratana, B.; Frey, P. A. “Dehydration is catalyzed by glutamate-136 and aspartic acid-135 active site residues in *Escherichia coli* dTDP-glucose 4,6-dehydratase” *Biochemistry* **2001**, 40, 12497.

129) Chen, Huawei ph.D dissertation 1998.

130) Beyer, Noelle ph.D dissertation, 2004.

131) Alam, Jenefer M.S. dissertation, 2003.

132) Sofia, H. J.; Chen, G.; Hetzler, B. G.; Reyes-Spindola, J. F.; Miller, N. E. “Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods” *Nucleic Acids Res.* **2001**, 29, 1097.

133) Ruzicka, F. J.; Lieder, K. W.; Frey, P. A. “Lysine 2,3-aminomutase from *Clostridium subterminale* SB4: mass spectral characterization of cyanogens bromide-treated peptides and cloning, sequencing, and expression of the gene *kam*A in *Escherichia coli*.” *J. Bacteriol*. **2000**, 182, 469.

134) Frey, P. A. “Radical reactions featuring lysine 2,3-aminomutase” *Comprehensive Chemistry of Natural Products* Amsterdam: Elsevier, 1999, vol 5, 205.

135) Zhao, Lishan ph.D dissertation, 2000.

136) Brennan, P. J.; Nikaido, H. “The envelope of mycobacteria” *Annu Rev Biochem*. **1995**, 64, 29.

137) Malthouse, J. P. “Stereospecificity of alpha-proton exchange reactions catalysed by pyridoxal-5'-phosphate-dependent enzymes” *Biochim. Biophys. Acta*. **2003**, 1647(1-2), 138.

138) Badet-Denisot, M. -A.; Rene, L.; Badet, B. “Mechanistic investigations on glucosamine-6-phosphate synthase” *Bull. Soc. Chim. Fr*. **1993**, 130, 249.

139) Obmolova, G.; Badet-Denisot, M. A.; Teplyakov, A. “Cyrstallization and preliminary X-ray analysis of the two domains of glucosamine-6-phosphate synthase from *Escherichia Coli*” *J. Mol. Bio*. **1994**, 242, 703.

140) Teplyakov, A.; Obmolova, G.; Badet-Denisot, M. A.; Badet, B.; Polikarpov, I. “Involvement of the C terminus in intramolecular nitrogen channeling in glucosamine 6-phosphate synthase: evidence from a 1.6 A crystal structure of the isomerase domain” *Structure* **1998**, 6, 1047.

141) Badet-Denisot, M. A.; Fernandez-Herrero, L. A.; Berenguer, J.; Ooi, T.; Badet, B. “Characterization of L-glutamine: D-fructose-6-phosphate amidotransferase from an extreme thermophile *Thermus Thermophilus* HB8” *Arch. Biochem. Biophys*. **1997**, 337, 129.

142) Golinelli-Pimpaneau, B.; Badet, B. “Possible involvement of Lys603 from *E. coli* glucosamine-6-phosphate synthase in the binding of its substrate fructose 6-phosphate” *Eur. J. Biochem*. **1991**, 201, 175.

143) Sparks, T. C.; Thompson, G. D.; Kirst, H. A.; Hertline, M. B.; Larson, L. L.; Worden, T. V.; Thibault, S. T. “Biological activity of the spinosyns, new fermentation derived insect control agents, on tobacco budworm larvae” *J. Econ. Entomol*. **1998**, 91, 1277.

144) Crouse, G. D.; Sparks, T. C. “Naturally derived materials as products and leads for insect control: the spinosyns” *Rev. Toxicol*. **1998**, 2, 133.

145) Sparks, T. C.; Thompson, G. D.; Kirst, H. A.; Hertlein, M. B.; Mynderse, J. S.; Turner, J. R.; Worden, T. V. “Fermentation-derived insect control agents: the spinosyns” *Methods Biotechnol*. **1999**, 5, 171.

146) Kirst, H. A.; Michel, K. H.; Mynderse, J. S.; Chio, E. H.; Yao, R. C.; Nakatsukasa, W. M.; Boeck, L.; Occolowitz, J. L.; Paschal, J. W.; Deeter, J. B.; Thompson, G. D. *Synthesis and Chemistry of Agrochemicals III*; Baker, D. R., Ed.; American Chemical Society: Washington, DC, 1992, pp 214-225.

147) Waldron, C.; Matsushima, P.; Rosteck, P. R. Jr.; Broughton, M. C.; Turner, J.; Madduri, K.; Crawford, K. P.; Merlo, D. J.; Baltz, R. H. “Cloning and analysis of the spinosad biosynthetic gene cluster of *Saccharopolyspora spinosa*” *Chem. Biol.* **2001**, 8, 487.

148) Matsushima, P.; Broughton, M. C.; Turner, J. R.; Baltz, R. H. “Conjugal transfer of cosmid DNA from *Escherichia coli* to *Saccharopolyspora spinosa*: effects of chromosomal insertions on macrolide production”*Gene* **1994**, 146, 39.

149) Laemmli, U. K. “Cleavage of structural proteins during the assembly of the head of bacteriophage T4” *Nature* **1970**, 227, 680.

150) Vesterberg, O. “Staining of protein zones after isoelectric focusing in polyacrylamide gels” *Biochim. Biophys. Acta*. **1971**, 243, 345.

151) Andrew, P. “Estimation of the molecular weights of proteins by Sephadex gel-filtration” *Biochem. J*. **1964**, 91, 222.

152) Moffatt, J. G.; Khorana, H. G. “Nucleotide polyphophates. X. The synthesis and some reactions of nucleoside 5’-phosphoromorpholidates and related compounds. Improved methods for the preparation of nucleoside 5’-polyphosphates” *J. Am. Chem. Soc*. **1961**, 83, 649.

153) Personal communication with Charles Evans Melancon.

154) Hanahan, D. *DNA cloning*; IRL Press: New York, 1985.

155) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular cloning, a laboratory manual*; 2nd ed.; Cold Spring Harbor Press: Cold Spring Harbor, NY, 1989.

156) Ogasawara, Y.; Katayama, K.; Minami, A.; Otsuka, M.; Eguchi, T.; Kakinuma, K. “Cloning, sequencing, and functional analysis of the biosynthetic gene cluster of macrolactam antibiotic vicenistatin in *Streptomyces halstedii*” *Chem. Biol*. **2004**, 11, 79.

157) Chatwood, L. L; Muller, J.; Gross, J. D.; Wagner, G.; Lippard, S. J. "NMR Structure of the Flavin Domain from Soluble Methane Monooxygenase Reductase from Methylococcus capsulatus (Bath)" *Biochemistry* **2004**, 43(38), 11983.

158) Fox, B. G.; Froland, W. A.; Dege, J. E.; Lipscomb, J. D. "Methane monooxygenase from *Methylosinus trichosporium* OB3b" *J. Biol. Chem*. **1989**, 264(17), 10023.

159) Barquera, B.; Nilges, M. J.; Morgan, J. E.; Ramirez-Silva, L.; Zhou, W.; Gennis, R. B. "Mutagenesis Study of the 2Fe-2S Center and the FAD Binding Site of the Na(+)-Translocating NADH:Ubiquinone Oxidoreductase from Vibrio cholerae" *Biochemistry* **2004**, 43(38), 12322.

160) Cafaro, V.; Scognamiglio, R.; Viggiani, A.; Izzo, V.; Passaro, I.; Notomista, E.; Dal-Piaz, F.; Amoresano, A.; Casbarra, A.; Pucci, P.; Donato, A. D. "Expression and purification of the recombinant subunits of toluene/*o*-xylene monooxygenase and reconstitution of the active complex" *Eur. J. Biochem*. **2002**, 269, 5689.

161) Fu, M.; Jack, R. F.; Morgan, T. V.; Dean, D. R.; Johnson, M. K. "*nifU* gene product from *Azotobacter vinelandii* is a homodimer that contains two identical [2Fe-2S] clusters" *Biochemistry* **1994**, 33, 13455.

162) Flamigni, F.; Guarnieri, C.; Caldarera, C. M. "Rat liver cytosol contains NADPH- and GSH-dependent factors able to restore ornithine decarboxylase inactivated by removal of thiol reducing agents" *Biochem. J*. **1988**, 250(1), 53.

163) Eto, K.; Kimura, H. “The production of hydrogen sulfide is regulated by testosterone and *S*-adenosyl-L-methionine in mouse brain” *J. Neurochem*. **2002**, 83(1), 80.

164) Linnebank, M.; Janosik, M.; Kozich, V.; Pronicka, E.; Kubalska, J.; Sokolova, J.; Linnebank, A.; Schmidt, E.; Leyendecker, C.; Klockgether, T.; Kraus, J. P.; Koch, H. G. “The cystathionine beta-synthase (CBS) mutation c.1224-2A>C in Central Europe: Vitamin B6 nonresponsiveness and a common ancestral haplotype” *Hum. Mutat*. **2004**, 24(4), 352.

165) Hyun, C.; Kim, S. S.; Sohng, J. K.; Hahn, J.; Kim, J.; Suh, J. “An efficient approach for cloning the dNDP-glucose synthase gene from actinomycetes and its application in Streptomyces spectabilis, a spectinomycin producer” *FEMS Microbiol. Lett*. **2000**, 183(1), 183.

166) August, P. R.; Tang, L.; Yoon, Y. J.; Ning, S.; Muller, R.; Yu, T. W.; Taylor, M.; Hoffmann, D.; Kim, C. G.; Zhang, X.; Hutchinson, C. R.; Floss, H, G. “Biosynthesis of the ansamycin antibiotic rifamycin: deductions from the molecular analysis of the rif biosynthetic gene cluster of Amycolatopsis mediterranei S699” *Chem. Biol*. **1998**, 5(2), 69.

167) Chen, H.; Zhao, Z.; Hallis, T. M.; Guo, Z.; Liu H. W. “Insights into the branched –chain formation of mycarose: methylation catalyzed by an (*S*)-adenosylmethionine-dependent methyltransferase” *Angew. Chem. Int. Ed*. **2001**, 40(3), 607.

168) Di Croce, L.; Buschbeck, M.; Gutierrez, A.; Joval, I.; Morey, L.; Villa, R.; Minucci, S. “Altered Epigenetic Signals in Human Disease” *Cancer Biol. Ther.* **2004**, 3(9).

169) Li, L, C.; Okino, S. T.; Dahiya, R. “DNA methylation in prostate cancer” *Biochim. Biophys. Acta*. **2004**, 1704(2), 87.

170) Carvin, C. D.; Parr, R. D.; Kladde, M. P. “Site-selective in vivo targeting of cytosine-5 DNA methylation by zinc-finger proteins” *Nucleic Acids Res*. **2003**, 31(22), 6493.

171) De Brasi, C. D.; Bowen, D. J.; Collins, P. W.; Larripa, I. B. “The CpG island in intron 22 of the factor VIII gene is predominantly methylated on the X chromosome of human males” *J. Hum. Genet*. **2002**,47(5), 239.

172) Trievel, R. C. “Structure and function of histone methyltransferases” *Crit. Rev. Eukaryot. Gene Expr*. **2004**, 14(3), 147.

173) Tsai, C C.; Fondell, J. D. “Nuclear receptor recruitment of histone-modifying enzymes to target gene promoters” *Vitam. Horm*. **2004**, 68, 93.

174) Duerre, J A.; Onisk, D. V. “Specificity of the histone lysine methyltransferases from rat brain chromatin” *Biochim. Biophys. Acta*. **1985**, 843(1-2), 58.

175) Ma, H.; Baumann, C. T.; Li, H.; Strahl, B. D.; Rice, R.; Jelinek, M. A.; Aswad, D. W.; Allis, C. D.; Hager, G. L.; Stallcup, M. R. “Hormone-dependent, CARM1-directed, arginine-specific methylation of histone H3 on a steroid-regulated promoter” *Curr. Biol*. **2001**, 11(24), 1981.

176) Poehlsgaard, J.; Douthwaite, S. “Macrolide antibiotic interaction and resistance on the bacterial ribosome” *Curr. Opin. Investig. Drugs* **2003**, 4(2), 140.

177) Miyamoto, Y.; Johdo, O.; Nagamatsu, Y.; Yoshimoto, A. “Cloning and characterization of a glycosyltransferase gene involved in the biosynthesis of anthracycline antibiotic beta-rhodomycin from *Streptomyces violaceus*” *FEMS Microbiol. Lett*. **2002**, 206(2), 163.

178) Wu, G.; Mansy, S. S.; Wu, Sp. S. P.; Surerus, K. K.; Foster, M. W.; Cowan, J. A. “Characterization of an iron-sulfur cluster assembly protein (ISU1) from *Schizosaccharomyces pombe*” *Biochemistry* **2002**, 41(15), 5024.

179) Muhlenhoff, U.; Richhardt, N.; Gerber, J.; Lill, R. “Characterization of iron-sulfur protein assembly in isolated mitochondria. A requirement for ATP, NADH, and reduced iron” *Biol. Chem*. **2002**, 277(33), 29810.

180) Inano, H.; Tamaoki, B. “Purification of NADPH-cytochrome P-450 reductase from microsomal fraction of rat testes, and its chemical modification by tetranitromethane” *J. Steroid. Biochem*. **1986**, 25(1), 21.

181) Oppermann, F. B.; Schmidt, B.; Steinbuchel, A. “Purification and characterization of acetoin:2,6-dichlorophenolindophenol oxidoreductase, dihydrolipoamide dehydrogenase, and dihydrolipoamide acetyltransferase of the Pelobacter carbinolicus acetoin dehydrogenase enzyme system” *J. Bacteriol*. **1991**, 173(2), 757.

182) Aliverti, A.; Piubelli, L.; Zanetti, G.; Lubberstedt, T.; Herrmann, R. G.; Curti, B. "The role of cysteine residue of spinach ferredoxin-NADP$^+$ reductase as assessed by site-directed mutagenesis" *Biochemistry* **1993**, 32(25), 6374.

183) Frey, P. A. "Radical mechanisms of enzymatic catalysis" *Annu. Rev. Biochem*. **2001**, 70, 121.

184) Beinert, H.; Thomson, A. J. "Three-iron clusters in iron-sulfur proteins" *Arch. Biochem. Biophys*. **1983**, 222(2), 333.

185) Berkovitch, F.; Nicolet, Y.; Wan, J. T.; Jarrett, J. T.; Drennan, C. L. "Crystal structure of biotin synthase an *S*-adenosylmethionine-dependent radical enzyme" *Science* **2004**, 303, 76.

186) Murataliev, M B.; Feyereisen, R.; Walker, F. A. "Electron transfer by diflavin reductases" *Biochim. Biophys. Acta*. **2004**, 1698(1), 1.

187) Ceccarelli, E. A.; Arakaki, A. K.; Cortez, N.; Carrillo, N. "Functional plasticity and catalytic efficiency in plant and bacterial ferredoxin-NADP(H) reductases" *Biochim. Biophys. Acta*. **2004**, 1698(2), 155.

188) Carrillo, N.; Ceccarelli, E. A. "Open questions in ferredoxin-NADP+ reductase catalytic mechanism" *Eur. J. Biochem*. **2003**, 270(9), 1900.

189) Unpublished results from our laboratory.

190) Alessio, A. C.; Annichino-Bizzacchi, J. M.; Bydlowski, S. P.; Eberlin, M. N.; Vellasco, A. P.; Hoehr, N. F. "Polymorphisms in the methylenetetrahydrofolate reductase and methionine synthase reductase genes and homocysteine levels in Brazilian children" *Am. J. Med. Genet*. **2004**, 128A(3), 250.

191) Dorweiler, J. S.; Finke, R. G.; Matthews, R. G. “Cobalamin-dependent methionine synthase: probing the role of the axial base in catalysis of methyl transfer between methyltetrahydrofolate and exogenous cob(I)alamin or cob(I)inamide” *Biochemistry*. **2003**, 42(49), 14653.

192) Cosper, M. M.; Jameson, G. N. L.; Hernandez, H. L.; Krebs, C.; Huynh, B. H.; Johnson, M. K. “Characterization of the cofactor composition of *Escherichia coli* biotin synthase” *Biochemistry* **2004**, 43, 2007.

193) Bates, C. J.; Pentieva, K. D.; Matthews, N.; Macdonald, A. “A simple, sensitive and reproductive assay for pyridoxal 5’-phosphate and 4-pyridoxic acid in human plasma” *Clinica. Chimica. Acta*. **1999**, 280, 101.

194) Ollagnier-de-Choudens, S.; Mulliez, E.; Hewitson, K. S.; Fontecave, M. “Biotin synthase is a pyridoxal phosphate-dependent cysteine desulfurase” *Biochemistry* **2002**, 41, 9145.

195) Chen, H.; Yamase, H.; Murakami, K.; Chang, C. W.; Zhao, L.; Zhao, Z.; Liu, H. W. “Expression, purification, and characterization of two *N*,*N*-dimethyltransferases, TylM1 and DesVI, involved in the biosynthesis of mycaminose and desosamine” *Biochemistry* **2002**, 41, 9165.

VITA

Lin Hong was born January 17, 1976 in ShenYang, P. R. China to JiQiu Xu and QingLan Hong. After graduating from the ShenYang Second High School (ShenYang, P. R. China) in 1993, she attended University of Science and Technology of China in Hefei, P. R. China for her undergraduate studies. She received her B. A. in Chemistry in 1998. Later that year she began her graduate studies in Chemistry at the University of Minnesota-Twin Cities campus. Her interest in Bioorganic Chemistry and Enzymology motivated her to join the research group of Hung-wen (Ben) Liu. Lin completed a two-year teaching assistantship for General Chemistry. In the fall of 2000, Lin moved with Ben Liu and six other group members to continue her graduate studies at the University of Texas at Austin, Department of Chemistry and Biochemistry.

Permanent Address: 7201 Hart Ln, #2062, Austin, TX, 78731

This dissertation was typed by Lin Hong.